

UCLA

UCLA Electronic Theses and Dissertations

Title

Empirically Guided Coordination of Multiple Evidence-Based Treatments: Relevance Mapping

Permalink

<https://escholarship.org/uc/item/72m2v382>

Author

Bernstein, Adam

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Empirically Guided Coordination of Multiple Evidence-Based Treatments:

Relevance Mapping

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Psychology

by

Adam Drew Bernstein

2012

GENERAL ABSTRACT OF THE DISSERTATION

Empirically Guided Coordination of Multiple Evidence-Based Treatments:

Relevance Mapping

by

Adam Drew Bernstein

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2012

Professor Bruce Chorpita, Chair

Well-designed studies have produced over 300 effective treatments that have been summarized in numerous lists of evidence-based treatments (EBTs). At the same time, the field is making great gains in the understanding of how to implement those treatments, once chosen. However, there is no structured guidance for how to select an optimal set of EBTs from those lists that is maximally relevant and minimally redundant with respect to its fit for a targeted service sample. This dissertation introduces relevance mapping, a methodology that addresses this problem. This dissertation consists of three studies that respectively describe the methodology, and use it to evaluate two open questions regarding treatment coordination. Relevance mapping uses automated comparison of the characteristics of each child in a targeted service sample to the participant characteristics from every study of every successful treatment. Relevance mapping addresses who is and is not coverable by any EBT in the literature, under configurable assumptions about which features must match between study participants and children in the service sample. Relevance mapping can then identify the minimum set of

treatments needed to serve the maximum number of children in the service sample, based on those same user-defined matching features. The first study describes this methodology in detail along with the context of the problem it addresses within the framework of knowledge management in mental health. The second study compares the efficiency of relevance mapping results when treatments are defined as intact programs or as collections of their constituent procedures. Finally, the third study applies relevance mapping to a large mental health service agency sample to assess the degree to which EBTs fit the problems, demographics, and treatment settings of youths served using wraparound process. Wraparound is a widely implemented and highly popular model for organizing individualized treatments and supports for children with complex needs. However its effectiveness has long been in question, making the combination of wraparound and EBTs an intriguing possibility. The dissertation's overarching goal is to illustrate a methodology for better application of the evidence base to applied settings, under a variety of different definitions and assumptions.

The dissertation of Adam Drew Bernstein is approved.

Mary Jane Rotheram-Borus

Steven Reise

Steve Lee

Bruce Chorpita, Committee Chair

University of California, Los Angeles

2012

TABLE OF CONTENTS

| | | |
|------|---|----|
| I. | General Abstract of the Dissertation | ii |
| II. | Chapter 1: Empirically Guided Coordination of Multiple Evidence-Based Treatments: An Illustration of Relevance Mapping in Children's Mental Health Services | 1 |
| a. | Abstract | 2 |
| b. | Introduction | 3 |
| | i. Selecting Locally Relevant Sets of Treatments | 4 |
| | ii. A Missing Link in the “Flow of Knowledge” | 6 |
| | iii. Relevance Mapping | 9 |
| c. | Method | 10 |
| | i. Client Dataset | 10 |
| | ii. Additional Datasets | 12 |
| | iii. Procedure | 15 |
| | iv. Data Analysis | 18 |
| d. | Results | 19 |
| e. | Discussion | 22 |
| f. | Tables | 31 |
| g. | Figures | 33 |
| h. | References | 36 |
| III. | Chapter 2: How We Define Treatment Affects Coverage and Efficiency in Coordinating Evidence Based Practice | 41 |
| a. | Abstract | 42 |

| | | |
|------|---|----|
| b. | Introduction | 43 |
| i. | Approaches to Treatment Definition | 43 |
| ii. | Levels of Analysis of Treatment Operations | 45 |
| iii. | Effects of Treatment Definition on Coverage and Coordination | 47 |
| iv. | The Present Study | 29 |
| c. | Method | 49 |
| i. | Client Sample | 49 |
| ii. | Study Datasets | 50 |
| iii. | Procedure | 52 |
| iv. | Data Analysis | 54 |
| d. | Results | 56 |
| i. | Aim 1: Differences in coverage | 57 |
| ii. | Aim 2: Efficiency of programs and practice Elements | 58 |
| iii. | Aim 3: A hybrid model | 60 |
| e. | Discussion | 61 |
| f. | Tables | 70 |
| g. | References | 78 |
| IV. | Chapter 3: Investigating the Fit Between Youths Served by Wraparound Process and Evidence-Based Treatments | 83 |
| a. | Abstract | 84 |
| b. | Introduction | 86 |

| | |
|-----------------------|-----|
| i. Wraparound Process | 87 |
| ii. The Present Study | 90 |
| c. Method | 92 |
| i. Client Sample | 92 |
| ii. Study Datasets | 94 |
| iii. Procedures | 96 |
| iv. Data Analysis | 97 |
| d. Results | 99 |
| e. Discussion | 103 |
| f. Tables | 111 |
| g. References | 117 |

LIST OF FIGURES

| | |
|---|----|
| Chapter 1, Figure 1. An illustration of the relation of evidence and practice in the context of children's mental health services | 31 |
| Chapter 1, Figure 2. Graphical illustration of the relevance mapping algorithm for determining client coverage | 32 |

LIST OF TABLES

| | | |
|---------------------|--|-----|
| Chapter 1, Table 1. | Children not coverable (NC) by evidence-based treatments identified in published randomized clinical trials (N studies = 435), assuming the prior availability of Multisystemic Therapy (MST) for disruptive behavior | 31 |
| Chapter 1, Table 2. | Treatments relevant to youths in the PAG scenario with coverage criterion of 100% of coverable youth, assuming the prior availability of a Multisystemic Therapy for disruptive behavior. | 32 |
| Chapter 2, Table 1. | Study dataset characteristics. All studies in the study datasets were randomized clinical trials (RCTs). Protocols represent the actual manuals or other descriptions of treatment tested in an RCT, while a study group represents those participants in a study who received a specific treatment protocol. | 70 |
| Chapter 2, Table 2. | Children not coverable by evidence-based treatments identified in published randomized clinical trials (RCTs) corresponding to programs listed by the National Registry of Effective Practices and Programs (NREPP), the California Institute of Mental Health (CIMH), and by common practice elements (PEs) coded from all identified RCTs. | 71 |
| Chapter 2, Table 3. | Programs relevant to youths with coverage criterion of 100% of coverable youths | 72 |
| Chapter 2, Table 4. | Practice Elements (PEs) relevant to youths with coverage criterion of 100% of coverable youths. | 74 |
| Chapter 3, Table 1. | Youths diagnostic and demographic characteristics and percentage of children in the corresponding categories not coverable (NC) by evidence-based treatments identified in published randomized clinical trials (N studies = 524) | 111 |
| Chapter 3, Table 2. | Practice Elements (PEs) relevant to youths in the PAGS scenario with coverage criterion of 100% of coverable youth | 113 |
| Chapter 3, Table 3. | Practice Elements (PEs) relevant to youths in the PAGS scenario with coverage criterion of 98% of coverable | 115 |

youth

ACKNOWLEDGEMENTS

I would like to thank the following co-authors for their contributions to this work.

Specifically, Chapter One is a version of:

Chorpita, B. F., Bernstein, A., & Daleiden, E. L (in press). Empirically guided coordination of evidence-based treatment: An illustration of Relevance Mapping in children's mental health services. *Journal of Consulting and Clinical Psychology*.

Chapter Two is a version of:

Bernstein, A., Chorpita, B.F., Daleiden, E. L, Ebesutani, C, & Rosenblatt, A. *How we define treatment affects coverage and efficiency in coordinating evidence based practice*. Manuscript in preparation.

Chapter Three is a version of:

Bernstein, A., Chorpita, B.F., Rosenblatt, A., Becker, K. D., & Daleiden, E. L. *Investigating the Fit Between Youths Served by Wraparound Process and Evidence-Based Treatments*. Manuscript in preparation.

VITA

2004 B.S., Computer Science
Stanford University
Stanford, California

2004 M.A., Psychology
Stanford University
Stanford, California

2006-2008 Graduate Student Research Assistant
Department of Psychology
Honolulu, Hawaii

2008 Adjunct Faculty Instructor
University of Hawaii
Honolulu, Hawaii

2008-2010 Graduate Student Research Assistant
Department of Psychology
University of California, Los Angeles

2009 Graduate Summer Research Mentorship Program Fellowship
Department of Psychology
University of California, Los Angeles

2010 Teaching Assistant
Department of Psychology
University of California, Los Angeles

2010-2011 Graduate Research Mentorship Program Fellowship
Department of Psychology
University of California, Los Angeles

2011-2012 Predoctoral Psychology Intern
UCLA Semel Institute for
Neuroscience and Human Behavior

SELECTED PUBLICATIONS

Chorpita, B. F., Rotheram-Borus, M. J., Daleiden, E. L., Bernstein, A. D., Cromley, T., Swendeman, D., & Regan, J. (in press). The old solutions are the new problem: How do we better use what we already know about reducing the burden of mental illness? Perspectives on Psychological Science.

Ebesutani, C., Bernstein, A., Chorpita, B. F., & Weisz, J. R. (in press). A transportable assessment protocol for prescribing youth psychosocial treatments in real-world settings: Reducing assessment burden via self-report scales. *Psychological Assessment*.

Chorpita, B. F., Bernstein, A., & Daleiden, E. L. (2011). Empirically guided coordination of multiple evidence-based treatments: An illustration of relevance mapping in children's mental health services. *Journal of Consulting and Clinical Psychology*, 79(4), 470-480.

Ebesutani, C., Bernstein, A., Martinez, J. I., Chorpita, B. F., & Weisz, J. R. (2011). The Youth Self Report: Applicability and validity across younger and older youths. *Journal of Clinical Child and Adolescent Psychology*.

Ebesutani, C., Bernstein, A., Nakamura, B. J., Chorpita, B. F., Higa-McMillan, C., & Weisz, J. R. (2010). Concurrent validity of the Child Behavior Checklist DSM-oriented scales: Correspondence with DSM diagnoses and comparison to syndrome scales. *Journal of Psychopathology and Behavioral Assessment*.

Chorpita, B. F., Bernstein, A., Miranda, J. (In press). Creating public health policy. In *The Oxford Handbook of Clinical Psychology*.

Ebesutani, C., Bernstein, A., Nakamura, B. J., Chorpita, B. F., & Weisz, J. R. (2010). A psychometric analysis of the Revised Child Anxiety and Depression Scale - parent version in a clinical sample. *Journal of Abnormal Child Psychology*, 38, 249-260.

Nakamura, B. J., Ebesutani, C., Bernstein, A., & Chorpita, B. F. (2009). A psychometric analysis of the Child Behavior Checklist DSM-oriented scales. *Journal of Psychopathology and Behavioral Assessment*, 31 (3), 178-189.

Chorpita, B. F., Bernstein, A., Daleiden, E. L., & Research Network on Youth Mental Health, Chicago, IL, US. (2008). Driving with roadmaps and dashboards: Using information resources to structure the decision models in service organizations. *Administration and Policy in Mental Health and Mental Health Services Research*. Special Issue: Improving Mental Health Services, 35(1-2), 114-123.

CHAPTER 1:

Empirically Guided Coordination of Multiple Evidence-Based Treatments:
An Illustration of Relevance Mapping in Children's Mental Health Services

ABSTRACT

Objective: Despite substantial progress in the development and identification of psychosocial evidence-based treatments (EBTs) in mental health, there is minimal empirical guidance for selecting an optimal *set* of EBTs maximally applicable and generalizable to a chosen service sample. Relevance mapping is a proposed methodology that addresses this problem through structured comparison of client characteristics in a service sample to participant characteristics from studies of EBTs. *Method:* We demonstrate the feasibility of relevance mapping using data from 1,781 youths in a statewide mental health system and a study dataset including 437 randomized clinical trials. Relevance mapping (a) reveals who is “coverable” by any EBT, under different definitions of matches between study participants and clients, and (b) identifies minimum sets of treatments needed to serve maximum numbers of clients, across different levels of analysis for defining treatment operations. *Results:* In the illustration sample, all problems targeted by the study dataset review were fully coverable when matching only required clients to have the same problem as EBT study participants. At the other extreme, when matching also required age, gender, ethnicity, and setting, the percentage of non-coverable youths increased to 86% in this sample. Two minimal sets of only 8 EBTs were identified that, when added to the one EBT already in place in that system, covered 100% of coverable youths when matching required problem, age, and gender. *Conclusions:* This methodology offers promise for the empirically guided selection and coordination of EBTs, thereby addressing one aspect of the gap between knowledge and practice.

INTRODUCTION

In the area of mental health services policy and research, the past 20 years have been characterized by a period of increased attention to service quality, through the application of rigorous standards of scientific evidence (Chambless & Hollon, 1998; Elliott, 1998; Glass & Arnkoff, 1996; Kazdin, 1996, 1998; Kendall, 1999; Kendall & Chambless, 1998; Nathan & Gorman, 1998; Nathan, Stuart, & Dolan, 2000; VandenBos, 1996; Weisz, Hawley, Pilkonis, Woody, & Follette, 2000). *Evidence-based treatments (EBTs)* have become the central mechanism for the application of scientific findings to practice delivered in service settings, and as the number of EBTs has grown, numerous lists and catalogues have attempted to organize what is known about them (e.g., Substance Abuse and Mental Health Services Administration [SAMHSA], 2011). In models connecting health science to health care practice more generally (cf. Frenk, 2009), Graham et al. (2006) refer to this cataloguing phase as *knowledge synthesis*, i.e., “the application of explicit and reproducible methods to the identification, appraisal, and synthesis of studies or information relevant to specific questions.” (p. 19). In mental health, these reviews, meta-analyses, and registries serve as lists from which one can select a single EBT that is well-suited to a single client (e.g., an evidence-based treatment for a depressed teen).

Given a long-standing national investment in both the development of new treatments and cataloguing those that are evidence-based, there has followed an increase in emphasis on understanding implementation or installation of the practices on those lists (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Glisson et al., 2008), topics that have been emphasized in nearly every major national report or action agenda relevant to mental health in the past 10 years (e.g., Hogan, 2003; Institute of Medicine, 2001; National Advisory Mental Health Council Workgroup on Services Research and Clinical Epidemiology, 2006; SAMHSA, 2006). Increased

understanding of the specific challenges to implementing EBTs is a part of the strategic mission of the National Institute of Mental Health (2008) as well as the U.S. Department of Health and Human Services (2007).

Unfortunately, despite these considerable investments and initiatives over a nearly 20 year period, the connection between evidence and practice in health care has been inefficient and fragmented, with approximately a third of all health practice being inconsistent with scientific findings, and more than 20% either unnecessary or harmful (e.g., Agency for Health Research and Quality, 2001; Schuster, McGlynn, & Brook, 2005). The gulf between evidence and practice may be even more severe in mental health, with the majority of services delivered in usual care settings having little or no relation to practice supported by research (e.g., Zima et al., 2005). No single reason appears to account for this gap; rather, it is believed to stem from a multiplicity of challenges that can arise within a complex cycle or flow (e.g., Frenk, 2009; Graham et al., 2006) representing the connection between evidence and practice. Many of those challenges are well-known and have been written about for decades, such as practitioners' lack of access to clinical research, as well as limits to the applicability of many clinical research findings to clinical practice (e.g., Barlow, 1981; Shafran et al., 2009).

Selecting Locally Relevant Sets of Treatments

However, other obstacles are more insidious such as the challenge in selecting an optimal *array* of EBTs from a larger list or registry in order to serve a known population—a problem that, although subtle, is now emerging as a major challenge as large service organizations struggle to allocate training resources efficiently and to organize their workforce and service arrays to be consistent with policies emphasizing the use of EBTs. How, for example, should a school-based mental health system with limited fiscal and personnel resources know which *set* of

EBTs from a national list would serve the largest number of its students? Likewise, a county-funded system whose resources are already over-extended may have to eliminate two EBTs from its service array to deal with budget cuts. Dropping which two would impact the fewest clients? Because there is no formal analytic model to inform such questions, these decisions are now typically made with minimal empirical guidance, and often involve a combination of rational heuristics (e.g., cut the most expensive treatment program, add a treatment program that worked in a similar organization), consensus-based meetings (e.g., voting on which programs get added or dropped), and stakeholder nominations (e.g., multiple interested parties lobbying in favor of their single preferred treatment, with minimal consideration for how that treatment will complement or compete with other treatments in the ultimate service array). We are thus badly in need of a model for how to select multiple treatments based on what the research says best fits a service population. Ironically, this problem of how to select *sets* of treatments stems from our overabundance of knowledge—a proliferation of EBTs.

To address this problem, we have proposed a methodology—relevance mapping—that employs a structured comparison of clients from a service population with participants from the research studies in the evidence base. Our methodology was developed primarily with large practice organizations in mind (managed care plans, state or county systems), but it could also be used at the level of small practices or even single practitioners as well (although the computational effort might not be cost-effective on such a small scale).

A Missing Link in the “Flow of Knowledge”

The gap addressed by relevance mapping is illustrated in Figure 1, in which we have positioned a question mark at the juncture at which the flow of knowledge (i.e., the connection between evidence and practice; Frenk, 2009) can break down. The leftmost side of the figure

begins with a representation of “raw” evidence, which Graham et al. refer to as the *knowledge inquiry* phase of connecting evidence to practice (see also Haynes, 2001). In the context of clinical trial research, this phase is described as “the unmanageable multitude of primary studies or information of variable quality that is out there and that may or may not be easily accessed” (Graham et al., 2006, p. 18). The next phase, consolidation, involves the process of organizing that raw information to make it useful for decision making (as noted earlier, what is described by Graham et al. as *knowledge synthesis*). With respect to the research evidence in mental health treatment, such consolidation has taken several forms, including national and international reviews, such as those by the Cochrane collaborative or the National Registry of Effective Practices and Programs (NREPP; SAMHSA, 2011).

The next phase, coordination, involves a process of integrative reasoning from sources of consolidated knowledge (e.g., making decisions using lists of EBTs). Graham et al. (2006) refer to this phase in the cycle as involving *knowledge tools or products*, which should “provide explicit recommendations with the intent of influencing what stakeholders do” (p. 19). We argue that despite very sophisticated, extensive, and diverse efforts in the consolidation phase, the gap between evidence and practice cannot be fully bridged by service organizations until there are practical knowledge tools to guide specific critical decisions in the coordination phase of the flow—in this instance, how does a health system administrator or policymaker select from a list of recommended treatments to assemble an optimal array of treatments to best serve that system’s service population?

Once past this problem of coordination, implementation of those treatments in the ideal array can occur next. Fortunately, as is true with the consolidation phase, there have been great academic and federal funding emphases on understanding implementation or installation of new

practices in applied settings (Fixsen et al., 2005; Glisson et al., 2008). Although implementation research is still in arguably early stages, it is already well known for example that successful implementation of innovative programs or practices should include training, coaching, consultation, administrative supports, evaluation, and feedback (Fixsen et al., 2005). The final phase in Figure 1 represents the actual practice that might ensue as a result of the specific application of scientific findings. In a high-functioning knowledge-to-practice flow, these treatments would be informed by the best available consolidated evidence, coordinated to optimize the benefits of those treatments on the population being served, and implemented according to the best supported principles of implementation science.

Recent initiatives to connect mental health science to practice have effectively begun to address early phases of the problem (i.e., knowledge generation [randomized trials] and knowledge synthesis [evidence summary lists]), and research on implementation of EBTs is beginning to yield answers to a later phase of the problem (i.e., installation and implementation). However, we argue that less is known about the phase between those two: decision-making about the coordinated selection of an optimal array of treatments. In a world of fixed resources, a typical service organization must not only choose from among a proliferation of EBTs (e.g., the 173 programs listed by SAMHSA's National Registry as of January 2011; SAMHSA, 2011), but must also organize them in such a way as to maximize their collective impact on the intended service population. In a simplified example, a children's mental health system with the capacity to implement three EBTs would serve a greater number of youths by selecting EBTs that target three unrelated problems (e.g., anxiety, depression, and disruptive behavior, assuming those problems are well-represented), as opposed to selecting three similar EBTs for disruptive behavior only. The complexity of maximizing evidence-based coverage increases quickly when

one has to consider treatments that are relevant to a local service population based on more than just clinical problem or diagnostic focus (e.g., when one also must consider age, gender, and ethnicity).

This task of maximizing evidence-informed service coverage represents a classic set optimization problem (Collatz & Wetterling, 1975; Hromkovič, 2004), whose many parameters include various characteristics of the service population (e.g., diagnoses, age), definitions of effective practice, the nature of the research literature and the features of study participants from that literature, the impact of treating or not treating given individuals in the service population, workforce learning capacity and turnover, among many other variables. Service organizations, government systems, or solo practitioners wishing to proceed from lists of EBTs to the implementation of a subset of those EBTs (once chosen) would be well-served by an empirical model for their selection and coordination.

Relevance Mapping

We use the term *relevance mapping* to describe the analytic framework for approaching questions about the applicability of treatments or sets of treatments concurrently to a given service population. Although this framework could include a considerably high number of parameters along which to compare clients with research samples, we intend here only to illustrate the basic architecture of the model itself and to provide an initial demonstration. The central part of the model essentially involves a simulation of enrollment of clients in published research trials, through a structured comparison of elements from independent databases containing client characteristics, study/treatment characteristics, and workforce characteristics. The method is comparable to taking every client in a given service population and seeing which studies in the literature have participants with matching characteristics. These client

characteristics (i.e., the client dataset) can include demographic and clinical variables typically found in a health record, such as presenting problem(s), age, gender, ethnicity, or clinical service setting. Treatments that best fit those client characteristics are selected from corresponding trials in the research literature (i.e., the study dataset), whose membership is defined by the application of a particular standard of evidence, for example, “treatments tested in randomized clinical trials showing statistically significant group differences.” Finally, decisions about selection and coordination of treatments might account for the treatments that are already available through appropriately trained practitioners in a target organization (i.e., the workforce dataset). For example, analysis for a system that already has an EBT for childhood traumatic stress disorders would “residualize” the other data sets with respect to traumatic stress. Specifically, because the youths with traumatic stress are already known to be appropriately covered by an EBT, those youths can be removed from the client dataset, and the matching studies are removed from the study dataset. The analysis then proceeds to handle optimization of “everyone else,” who is not already covered by the existing treatment(s) available in the mental health service organization.

The primary aim of this study is to illustrate relevance mapping as a structured analytic strategy to help guide coordination of treatments, making use of client, study, and workforce datasets in the context of children’s mental health. The examples provided are not intended to be demonstrative of what works in general in children’s mental health, but rather are merely an illustration of the process of how evidence can be coordinated to fit a specific, local service population, under a given set of assumptions about what constitutes acceptable strength of evidence and about what parameters are reasonable to consider regarding treatment generalizability (e.g., age, gender, presenting problems).

METHOD

Client Sample

The sample in this report comes from the Child and Adolescent Mental Health Division (CAMHD) of the Hawaii Department of Health, a statewide mental health system for youths with intensive emotional and behavioral needs. This sample represents 1,781 youths registered with CAMHD and receiving services during the fiscal year 2007. For youths with multiple episodes of treatment with CAMHD, only the first episode was considered. The only inclusion criterion was data availability. To be included in the analysis, a child had to have data on at least one variable from among age, gender, ethnicity, primary problem or setting. Availability of one of these variables allows a child's data to be analyzed in at least one of the scenarios possible with these parameters. Otherwise, the applicability of the research findings to that particular child could not be estimated. Of the 1,781 youths in the sample, 1,151 were males (64.6%) and 630 were females (35.4%). Age of youths in the sample ranged from 1 to 19 years with a median age of 15 ($M=13.8$, $SD=3.2$). Youths' ethnic groups included Multiethnic ($n=993$; 55.8%), Caucasian ($n=219$; 12.3%), Native Hawaiian or Pacific Islander ($n=172$; 9.7%), Asian American ($n=129$; 7.2%), African American ($n=31$; 1.7%), Hispanic/Latino ($n=22$; 1.2%), and Native American/Alaskan ($n=10$; 0.6%), with ethnicity data unavailable for 205 youths (11.5%).

Chart diagnoses were based on Diagnostic and Statistical Manual of Mental Disorders (4th ed. [DSM-IV]; American Psychiatric Association, 2000) codes. Evaluations were performed by clinical staff, the Department of Education clinical staff, or contracted mental health care providers, and although state performance standards encourage the use of structured or semistructured interviews to arrive at a clinical diagnosis, no specific protocols were mandated. For the purposes of this investigation, DSM-IV diagnoses were cross-walked to 8 broad categories (the mapping of all diagnoses to problem groups are available upon request), shown in

Table 1, and frequencies are reported under the columns labeled “N.” Because only primary diagnosis was considered in this illustration, each youth is represented in only one category. The decision to use 8 broad problem categories was made to provide a common framework for matching clients to studies, given that research studies use a wide variety of taxonomies and methods for classifying and enrolling participants, not limited to diagnosis.

Additional Datasets

The study dataset involved codes from 437 randomized clinical trials of child mental health treatments corresponding to the following problem areas: anxiety (n studies = 125), attention/hyperactivity (n studies = 83), autism spectrum (n studies = 25), depression (n studies = 32), disruptive behavior (n studies = 192), eating disorders (n studies = 11), substance use (n studies = 18), and traumatic stress (n studies = 13); characteristics of this dataset have been reported in previous research (Chorpita et al., in press). Each study had numerous records in the dataset, each representing a single characteristic of participants included in the study (e.g., problems, ages, genders, and ethnicities) as well as the setting(s) in which treatment was provided, the name and type of treatment protocol used, and other study information not used in the current analysis (e.g., treatment format, therapist education level, etc.). Eight hundred and thirty two coded treatments were grouped into 98 general “families” of approaches (e.g., “Cognitive Behavior Therapy,” “Client Centered Therapy,” “Family Systems Therapy,” “Parent Management Training,” “Multisystemic Therapy,” etc.), consistent with the organization of traditional meta-analytic reviews (e.g., Weisz, Hawley, & Jensen-Doss, 2004).

In order to determine which of these treatments are candidates for analysis, relevance mapping requires the user to select or define some standard of evidence that identifies which treatments in the study dataset are considered evidence-based, (although it does not depend on

any single definition). In other words, any list of EBTs could be used for analysis (e.g., NREPP [SAMHSA, 2011], American Psychological Association’s Division 53 review of evidence-based practices [Silverman & Hinshaw, 2008]), and each list’s standard of evidence could theoretically produce a slightly different solution. For this illustration, we employed a standard of evidence that has been used over several years with this same service population, which is based largely on the criteria developed and employed by the Task Force on Promotion and Dissemination of Psychological Procedures (1995). According to this standard, a manualized treatment must (a) show statistically superior outcomes to a waitlist or no-treatment control group in at least two randomized trials, (b) show statistically superior outcomes to an active treatment or psychological placebo in at least one randomized trial, or (c) show equivalent outcomes to an already established evidence-based treatment in at least one randomized trial in which the average group size is at least 30 participants (see Chorpita et al., in press, for additional details). Again, this particular definition of evidence is not integral to the relevance mapping methodology, and any other rationally chosen standard of evidence could apply as well.

The problem areas covered by each randomized clinical trial in the study dataset were reduced to the same 8 broad categories used with the client dataset (shown in Table 1). This small number of broad problem categories was chosen after considering alternate problem definitions including specific DMS-IV diagnoses (more than 300 categories; e.g., “bipolar II disorder,” “anorexia nervosa”) and a reduced set of 36 diagnostic groupings (e.g., “mood disorders,” “eating disorders”). Because only 177 (40.8%) of the randomized trials in our review reported clinical diagnosis, there is a tradeoff between the precision with which a problem is described and the availability of data to inform a match. Thus, although inferences about youths with “major depressive disorder, single episode” (a diagnosis) versus “depressed mood” (one of

our 8 problem areas) can perhaps be made with greater confidence, the literature relevant to those clients, and hence the EBTs available for analysis, is greatly reduced in this tradeoff. That said, our decision here is simply to illustrate the relevance mapping approach, not to propose the best way to conduct a relevance mapping analysis, and clearly a choice to match on diagnostic categories would be rationally defensible and just as easily handled within this analytic framework. For illustration purposes, the choice is somewhat arbitrary.

The workforce dataset is a list of treatments that are already in use at an organization. As noted above, if those treatments already cover youths in the client data set, those clients are removed and considered coverable prior to the formal relevance mapping analysis. The workforce dataset for this illustration contained only a single treatment, Multisystemic Therapy (MST; Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 1998), a home- and community-based treatment for adolescent delinquency, which was already in place in the Hawaii CAMHD system. The study dataset (described above) contained 15 studies of MST, and together these studies covered youths of both genders, ages 10 to 17, with ethnicities including Asian American, African American, Caucasian, Hispanic/Latino, and Multiethnic, and settings of treatment including community-, home-, hospital-, and school-based care. Thus to be covered by the workforce dataset, a client had to match participants on the scenario parameters in *at least one* successful study of MST (e.g., a child between the ages of 10 and 17 who also had disruptive behavior as the primary problem, in a problem-age scenario). Youths in the CAMHD system whose characteristics indicated that MST was an appropriate treatment were removed from each analysis as the model dictated.

Procedure

Figure 2 depicts our implementation of the relevance mapping model. The figure shows that the study dataset in this example is organized around the *study group*, which is the unit to which specific treatment protocols can be linked. A study group represents those participants in a study who received a specific treatment protocol (e.g., a treatment or control group in a randomized trial). The characteristics of the study group participants thus determine with whom each treatment has been shown to work.

Defining treatments. Although a specific treatment protocol is directly linked to each study group (i.e., the manual used in the study to treat that group), we can also link treatment protocols to many other levels of analysis for defining treatment (Figure 2, *practice units*). A user may wish to conceptualize treatments as clusters of theoretically related manuals (e.g., family therapy, cognitive behavioral therapy, etc.), or to divide manuals into components of treatments (e.g., relaxation procedures, use of rewards, etc.). *Practice unit* is thus our generic term for a classification scheme of treatment activities at any level of analysis. Mapping to these units allows the relevance mapping procedure to find solutions (i.e., best sets of treatments) in the “language” most meaningful for a particular organization. For example, a user could consider multiple versions/editions of a given anxiety protocol tested in separate studies as a single treatment, and thus, a single practice unit (e.g., the multiple editions of the “Coping Cat” protocol are members of the practice unit, “Coping Cat”). Alternatively, a user could consider all evidence-based CBT protocols for anxiety as a single practice unit (the practice unit, “CBT for Anxiety”). In either case, relevance mapping would allow the user to find best sets of treatments using the corresponding definition (matching youths to “Coping Cat” versus to “CBT for Anxiety”). Again, relevance mapping does not depend upon any assumptions about what constitutes the *definitive* practice unit (e.g., manualized treatment programs vs. general treatment

types; cf. Rogers & Vismara, 2008), but instead will accept any level of analysis chosen by the user. Much like the issue with matching on diagnosis or broad problem labels, this level of analysis issue is arbitrary with respect to an illustration of the model. Questions about how psychosocial treatment operations are best defined for the purposes of consolidation, coordination, or implementation are explicitly not part of the scope of this paper. As noted above, we grouped treatments into 98 broad groupings or families (e.g., “Cognitive Behavior Therapy”) as practice units for this illustration.

Scenarios for matching. Relevance mapping analyses can be performed under a variety of scenarios, which represent combinations of rationally selected parameters from the study dataset and the client dataset, on which clients and research participants must match. For example, one might define an EBT as relevant to those clients with the same primary problem and within the same age range as the study in which that treatment was successfully tested, thus creating a “problem-age” scenario. Each member of the client sample is thus compared to each study group in the literature that has produced an EBT, and if there is a match on all parameters for that scenario (e.g., the study included participants with the same primary problem and with the same age), that study group by client combination is written as a record to a “match” dataset, because there is an instance of an EBT that matches the given member of the client sample (note that this requires data to be available both in the client record and in the study).

Because there is no restriction that clients (or study participants) have only one value per factor on which to match, relevance mapping can take into account complex variations of otherwise related scenarios. For example, matching on “problem” could mean a client has (1) the same primary problem as the primary problem of the study participants, (2) any problem (from among several) in common with any problems of the study participants, (3) the same primary

problem for which study participants were included in the study, and none of the problems for which study participants were excluded from the study, and so forth. In other words, the parameters for matching—for problem(s) or for any other variable of interest—are multivariate and thus can be configured to represent typical study inclusion and exclusion criteria.

Reports. Once the dataset of client to study group matches is developed, it is then used to produce a variety of reports. A primary question to be addressed involves the number of coverable members of the client dataset for a given scenario, i.e., for whom in the service sample is there *any* relevant EBT? A client is thus considered coverable if there is at least one EBT in the literature that produces a match for that client on all parameters of the scenario in question. Non-coverable clients are those who have no EBT that would generalize to them under a given scenario, using a given standard of defining an EBT (with higher standards leading to more non-coverable clients). We describe these *client reports* as those which refer to clients as dependent variables (e.g., % of clients coverable), and these reports can expose who in a given service population is or is not coverable under a variety of scenarios that specify requirements for considering a given treatment relevant to a given client.

We define *practice reports* as those that refer to practice units as dependent variables. The primary questions from this perspective are how many and which practice units are required to cover a criterion percentage of clients. Practice reports can compare the smallest practice groupings required to cover a given percentage of clients across various scenarios. For example, a report could compare the smallest set of treatments needed to achieve 95% coverage in a scenario requiring matching on problem and age to the smallest set needed to cover the same percentage when treatment setting is additionally required for matching. Scenarios requiring

matching on a greater number of parameters (e.g., age, gender) lead to more practice units (e.g., treatments, components of treatments) being needed to cover the same number of clients.

The process used to find the best sets of practice units is conceptually simple—albeit computationally intensive: test all possible combinations of treatments (at units defined by the user) and report the smallest combination(s) that meet(s) the desired coverage threshold. Testing a given grouping entails checking whether it corresponds to study groups in the literature that were found to match clients under the assumptions of a specific scenario. If a set of practice units matches enough clients, it is an available solution, and if it is among the smallest of the available solutions it can be flagged as one of the minimal or “optimal” sets. Because the number of possible practice arrays to test can be very large (it grows more than exponentially with the total number of practice units considered), this process is done via computer automation.

Data Analysis

Three scenarios are represented in the illustration’s analysis: Problem (P), Problem-Age-Gender (PAG), and Problem-Age-Gender-Ethnicity-Setting (PAGES). In this illustration, we configured matching for the Problem factor using youths’ primary problems only (i.e., those derived from their primary DSM-IV diagnoses), and ignoring (for the purposes of matching) their comorbid problems (e.g, for a client to match a treatment for anxiety, that client could have any comorbid problems/disorders, but must have a primary anxiety problem/disorder). Because MST was already in the CAMHD practice array, we accounted for its presence by creating a residual class for each scenario (i.e., clients not already coverable by MST). Youth coverability was analyzed under each of the three scenarios using this corresponding residual class of youths not coverable by MST. The initial youth coverability results were then used to identify a working scenario affording reasonable coverage so as to proceed with the identification of minimum

complimentary treatment sets. As noted above, practice units were based on grouping treatments into 98 families of theoretically related protocols sharing similar clinical strategies (e.g., parent management training, family therapy, cognitive behavioral therapy, etc.), comparable to clusters used in major lists of effective treatments (e.g., Society of Clinical Child and Adolescent Psychology, 2011; APA Presidential Task Force on Evidence-Based Practice, 2006).

RESULTS

Table 1 represents a sample *client report* examining the Hawaii CAMHD dataset. As can be seen in column P (in which matching was required on primary problem only), all problems that were the target of the study dataset review were fully coverable for that scenario. All youths with disruptive behavior (i.e., coverable by MST) were not in the residual sample, and thus were not analyzed. In column PAG, abrupt increases in the percentage of non-coverable youths with attention problems or autism spectrum problems were noted, presumably owing to their age, gender, or combination thereof (reasons for non-coverability can easily be identified by changing only one parameter at a time across scenarios). In the PAGES scenario, 1,538 of 1,561 youths in the residual sample (99%) were not coverable; however, because the full CAMHD sample included 220 youths already coverable by MST, 86% (1,538 of 1,781) of the full CAMHD sample was not coverable. Thus, with the exception of MST, this evidence base did not generalize well to this client sample under our strictest set of assumptions in this illustration. The size of the residual sample increased as the generalization scenarios become more demanding, given that the single treatment in the workforce dataset (MST) generalized to fewer of the full 1,781 youths under more demanding requirements. Although the 15 studies of MST in the study dataset together covered ethnicities matching 78.2% of the sample and treatment settings matching 68.5% of the sample, the PAGES scenario requirements of matching on ethnicity and

setting resulted in 392 (64.1%) youths with disruptive behavior found not coverable by MST, as opposed to just 57 (9.3%) in the PAG scenario. This increase is primarily attributable to the fact that often no *single study* of MST matched youths on both setting and ethnicity. For example, though the 12 studies of MST together covered both multiethnic youths and hospital-based treatment, no single study provided treatment in the hospital setting and included multiethnic youths, and *that combination* was thus not coverable by MST in this analysis.

A practice report for the client dataset in the “Problem Age Gender” scenario (Table 2) shows that it would be possible to serve the full 58% of “coverable” youths using 8 treatment types (plus MST covering the additional 555 youths removed from the residual sample). The first column shows that Cognitive Behavior Therapy (CBT) alone is an EBT applicable to 43% of clients (of the possible 58% to whom any EBT applies). The rightmost column refers to the percentage of youths who would no longer be coverable, were a particular approach to be dropped from the identified service array. The table thus shows that although CBT applied to 43% of the clients, only 9% were uniquely coverable by CBT.

CBT applied to this large portion of the sample in part because of the broad treatment groupings (“families” of approaches) used for this illustration. The CBT family included EBTs for anxiety, autism spectrum, depression, disruptive behavior, eating disorders, substance use, and traumatic stress, and CBT thus covered some youths from each of these problem areas. Other treatments listed in Table 2 provide complementary (though often overlapping) coverage. Parent Management Training (PMT) addresses disruptive behavior and attention/hyperactivity problems for youths with ages ranging from 2 to 15 across multiple studies. A closely related treatment family, PMT and Problem Solving covered youths in the same problem areas but with an age range of 0 to 13. Intensive Communication Training (ICT) and Intensive Behavioral

Treatment (IBT) are approaches that address autism spectrum problems. Again, both appeared in the solution sets because their efficacy has been demonstrated with slightly non-overlapping age ranges (ages 1 to 10 for ICT and ages 2 to 12 for IBT). Finally, Self Verbalization was present in one of the optimal treatment sets identified, and Social Skills and Medication was present in the other. Both of these treatment families had studies in the study dataset that met the standard of evidence for attention/hyperactivity problems with similar age ranges, though only Social Skills and Medication had demonstrated efficacy for males, whereas Self Verbalization covered both genders. Table 2 shows that these two treatment families applied to somewhat different percentages of youths overall, but each uniquely covered 1.8%. Since the uniquely covered youths were identical (in identity as well as percentage), either of the two families could be selected to complete an optimal set.

DISCUSSION

The most striking implication of these findings for practice organizations involves the efficiency of informing the treatment array design decision. Given that we identified 98 treatment types in the study dataset, the CAMHD system faced choosing from among an almost incalculable number of possible sets of treatments (e.g., even knowing in advance that exactly 8 treatments must be added to MST yields over 157 billion combinations to choose from), and yet the relevance mapping procedure allowed us to narrow the options precisely to those eight that applied to the maximum amount of coverable youths under various different assumptions. For organizations facing choices about how to select a limited number of maximally relevant treatments from among the growing list of EBTs, we do not imagine this task can be accurately performed without computational supports such as the methods we have outlined.

Beyond these immediate implications, inspection of the patterns of results can yield useful information about the specific fit of a service array to a population. The attention/hyperactivity problem area provides a useful example: although all youths are covered in the P scenario, almost half are outside the tested range of any EBT when matching on age and gender is additionally required (although a problem-age scenario is not shown in Table 1 to distinguish the effects of age and gender, a problem-gender scenario did not differ from a problem only scenario for attention/hyperactivity, and thus it is the age requirement causing this decrease in coverage). In line with those observations, the study dataset did not contain any evidence supporting psychosocial or combined treatments for attention problems for youths above age 13. The uncoverable youths with attention/hyperactivity in this sample are above this age, and so treating them could require “going off the evidence base,” if one considers age to be a requirement for treatment generalization. In such situations, best practice may be to adapt or extend the practices found to be generalizable under less stringent requirements. For example, the treatments found to be generalizable in the P scenario could be adjusted to be age appropriate for older youths (i.e., *adaptation*) or they could be used unmodified with this new population (i.e., *extension*). The literature is replete with such gaps with respect to certain parameters in some samples (e.g., ethnicity), and these reports provide an explicit index of for which youths treatment adaptations or extensions might be appropriate.

The PAGES scenario results in Table 1 also demonstrate how one can identify the gaps that most affect a given population by moving from one column to the next in this client coverage report. As with age and gender, the ethnicity and setting matching requirements were added at the same time in this demonstration and so their effects cannot be distinguished in Table 1. However, it is readily apparent that in combination these additional requirements for matching

youths to treatments leave almost no youths covered by an EBT. The CAMHD service population may be particularly ill-suited for generalizing existing EBTs with regard to ethnicity and setting because the youths are highly ethnically diverse (whereas most EBTs have been tested with more ethnically homogenous populations) and many receive intensive services in residential or community-based settings where EBTs are less often tested. Notably, even under these strict assumptions MST was still relevant to more than a third of the youths with disruptive behavior as their primary problem.

The setting parameter illustrates a third option for how to proceed when youths are not found coverable: if possible, change a non-covered youth's value on an offending matching parameter. In the setting example, in addition to adapting or extending of treatments from less restrictive scenarios, one could also consider *redirecting* a youth to a service setting where an otherwise matching EBT is available. For example, if no EBT is found in the PAGES scenario for a youth receiving treatment in residential care, a conceivable alternative is to redirect that youth to receive treatment in a home or community-based setting using an EBT that meets all the other PAGES requirements for that youth.

Another result of note is that the majority of the youths coverable were coverable by multiple treatment types. That is, dropping a treatment type that applied to a given percentage of youths left only a smaller percentage of youths uncovered by an EBT, as shown in Table 2. Intensive Behavioral Treatment, for example, was in the final solution due to its unique coverage of a small number of clients with autism spectrum diagnoses who were not covered by Intensive Communication Training. Covering these uniquely coverable youths requires the inclusion of treatments in the final solution that may apply to a very small percentage of the overall service

population; thus, approaches aimed at covering fewer than 100% of coverable clients are likely to yield dramatically smaller arrays of practice units (e.g., treatments) as solutions.

Although such findings are interesting in their own right, the larger implication is that relevance mapping appears to offer an organized empirical model or framework to inform treatment selection and coordination decisions, which at present constitute a significant gap between knowledge synthesis and EBT implementation. This model can enable an organization or the field at large to delineate more clearly the edges of our understanding by making visible where inferential leaps in generalizing treatments are large or small. Similarly, relevance mapping makes salient the situations where the desired standard of evidence is simply unavailable and thus adaptation, extension, or redirection is required. By exposing the scenarios in which youths are and are not coverable, relevance mapping reports can provide instrumental guidance regarding what evidence to fall back on when the highest standard of evidence is unavailable. For example, if no appropriate treatment is available for a given youth under the PAGES scenario, one might adapt or extend an EBT identified in the PAG scenario or redirect the youth to a treatment setting where an otherwise matching EBT has been shown to work. This methodology is extremely flexible in that it is not dependent upon specific definitions of treatment units (e.g., practices components vs. treatment types vs. brand-name EBT programs), definitions of matching parameters (so long as a common definition of these parameters is applied both to studies and to clients), or even standards defining what constitutes an EBT. All of these definitions can be configured prior to analysis, such that the method can be applied in many different practice and policy contexts.

Various extensions of this approach are also possible. For example, an extension of the practice minimization approach involves ensuring that identified treatment sets provide multiple

options relevant to each client. In our example we found the smallest solutions that required at least one treatment for each client; however organizations with sufficient resources may desire multiple EBT options for each client, and so for example, one could compute solutions that outline a service array with at least two EBT options per youth. Another extension involves examining coverage when “resource constraints” are introduced that place an upper limit on the number of practice units in a solution. For example, one might decide *a priori* that a given service system can only introduce three new treatments in a given time period, and thus, the analysis can be aimed at outlining a service array that maximizes the percentage of youths impacted by the introduction of only these three new treatments. In other words, if my agency can learn only three treatments this year, which three should they be?

Around this core methodology of structured comparisons of clients to study participants, there are obviously many parameters that can be manipulated to create simulations that map the relevance of the literature to a service sample, and we believe that examining variations on these computational exercises is likely to shed light on how best to select and coordinate EBTs to maximize implementation efforts. These different configurations ought to be examined in future research with a variety of different client samples to illustrate important treatment selection principles (those robust across assumptions) as well as implications for workforce development and implementation demand (how many treatments are needed depending on different starting assumptions?). Further, although inherently “local” in its approach, this methodology could also be applied to large, nationally representative client samples so as to illustrate general service arrays likely to be efficient on a large scale, reducing the need for each service community to perform its own relevance mapping analysis.

One broader implication is that EBTs that are otherwise equivalent regarding their efficacy may be considerably different in terms of their importance for local implementation. That is, the definition of what constitutes a relevant EBT for a given service system is not only a matter of the strength of the research evidence (although we contend that criterion remains a minimum condition), but also a matter of (1) assumptions about the importance of generalization parameters, and (2) the local context (the nature of the workforce and clients in that system). Thus, these methods show that knowing the relevance of a treatment to a specific clinical population, in the context of the other treatments in that system, is an important aspect of building a comprehensive service array in a local context. Although in the current illustration, the generalization parameters are configurable by the user (i.e., one can choose “problem-age” instead of “problem-age-gender-ethnicity” as a condition of the solution), the importance of these parameters for generalization will ultimately be addressed by studies that define these boundaries through successes and failures along specific parameters. In other words, if treatment A is known to work with girls, and explicitly known not to work with boys (as opposed to unresearched with boys), then gender is a necessary parameter in relevance mapping solutions involving the applicability of treatment A. Currently, empirical knowledge of generalization boundaries remains notably underdeveloped in the literature, and thus, decisions about generalizability parameters are for now likely best guided by local judgment.

One limitation of this model (or more precisely, its current illustration) is the notion that matching clients to studies examines study parameters as concurrent sets, as opposed to all combinations of those matching parameters. In other words, in our example, a study is considered relevant to an 8-year old African American boy if the study contained boys, 8-year olds, and African Americans. However, it is theoretically possible that the “matching” study did

not include a single 8-year old African American boy, or that if such a boy was present in the study that he was not among the treatment responders. This limitation is potentially addressable, but not without a study data set that includes client-level data from those studies (as opposed to the traditional frequency counts, means, or ranges used to describe an entire sample in a method section), an approach that has low feasibility at present. Given this limitation, analyses that base inferences about matching upon these aggregated parameters are inherently biased to over-identify matches. Thus, in this illustration, estimates of percentage coverable within scenario represent upper limits, and true values are likely to be somewhat lower.

Similarly, users of this approach must consider defining the proportion of matching cases within a study data set. For example, having *any* girls (i.e., at least one) in a study count as a match for a girl in the client dataset is a more liberal rule than requiring a minimum percentage of girls (e.g., > 30%) within that study. Unlike the matching combination problem, this issue of the proportion of matching study cases is more easily addressable within the model by defining matches as requiring a higher proportion of the matching parameter within the study group. Again, our current illustration used the most liberal approach to defining matches and thus likely overestimates the percentage of youths coverable across the various scenarios with this client dataset. A third domain where our illustration uses a liberal approach is with regard to capacity issues. Specifically, for these analyses, we assumed that if a given treatment were made available in a service organization, the organization would then have the capacity to serve an unlimited number of youths with that treatment. Future investigations will need to incorporate the complexities of expected utilization rates, provider caseload maxima, and provider learning capacities.

Another potential limitation with this model involves the difficulty of defining goals for a given service system. We need to be explicit that in this illustration we are making assumptions that “optimal” means maximizing coverage, but estimating “maximal expected impact” on a service population could involve a combination of percentage of youths covered, expected effects of each EBT (i.e., predicted effect sizes), and effects of usual care (i.e., observed effect sizes). Although in some ways, this complexity is a limitation, it is not a limitation inherent in the relevance mapping methodology. Decisions regarding whether a small benefit for many is superior to a large benefit for few, for example, are not addressable through computation. Rather, they are a matter of local preferences and values. Thus, the present illustration of relevance mapping is a preliminary example of the line of reasoning and analysis that one could pursue in the “coordination” domain in Fig. 1, and solutions that model expected effects are also possible once those user preferences are known.

Finally, a challenge to achieving wide use of the relevance mapping model is the need for access to specific, structured data sources. Regarding the study dataset, a structured database is needed, ideally containing comprehensive information about treatments and the characteristics of populations with whom they have been shown to be effective. In our illustration, we used a large (N = 437 studies) privately compiled database for this purpose. However, it should be noted that public consolidated lists of EBTs could potentially be used as well. For example, in SAMHSA’s National Registry (SAMHSA, 2011) each program is listed along with structured population characteristics for each of the PAGES parameters. With this information, a relevance-mapping-like analysis could be performed, bypassing the linking to study groups shown in Figure 2 and instead directly comparing rationally compiled treatment descriptions with clients. As a downside, such analysis would be restricted to the particular definitions and assumptions implicit

in that registry, rendering untestable the effects of those definitions and assumptions on the results. Regarding the client data, we used a large dataset (N = 1,781 youths) from a large state-run service organization, but any coded dataset from a service organization that contains matching parameters of interest would be suitable. Although relevance mapping may be most applicable to organizations that serve large populations, large client sample size is not essential and only affects relevance mapping results inasmuch as it allows better estimation of population prevalence of the parameters used for matching (e.g., how much does this year's caseload of 40% girls mean that next year's caseload will have 40% girls?).

In summary, we propose that this methodology and its possible extensions may be highly useful in conjunction with existing consolidated lists of EBTs. By examining the fit of treatments to local service populations, this methodology can begin to address the problems faced by organizations seeking to select treatments from a diverse and ever-growing array of options, thereby shifting the focus from solely identifying treatments that work to identifying *relevant* treatments that work, and even more importantly, identifying *sets of relevant treatments* that best work together.

Table 1

Children not coverable (NC) by evidence-based treatments identified in published randomized clinical trials (N studies = 435), assuming the prior availability of Multisystemic Therapy (MST) for disruptive behavior

| Problem type | Scenario | | | | | |
|---|----------|------|----------|------|----------|------|
| | P | | PAG | | PAGES | |
| | <i>N</i> | % NC | <i>N</i> | % NC | <i>N</i> | % NC |
| Disruptive behavior (all) | 612 | 0% | 612 | 0% | 612 | 64% |
| Already coverable by MST | 612 | 0% | 555 | 0% | 220 | 0% |
| Not already coverable by MST | 0 | N/A | 57 | 2% | 392 | 100% |
| Depression | 268 | 0% | 268 | 0% | 268 | 97% |
| Attention/hyperactivity | 264 | 0% | 264 | 48% | 264 | 98% |
| Traumatic stress | 125 | 0% | 125 | 2% | 125 | 99% |
| Substance use | 68 | 0% | 68 | 0% | 68 | 100% |
| Anxiety | 54 | 0% | 54 | 0% | 54 | 87% |
| Autism spectrum | 19 | 0% | 19 | 42% | 19 | 100% |
| Eating disorders | 1 | 0% | 1 | 0% | 1 | 100% |
| Other/missing | 370 | 100% | 370 | 100% | 370 | 100% |
| Total (not already coverable by MST) | 1,169 | 32% | 1,226 | 42% | 1,561 | 99% |
| Total | 1,781 | 21% | 1,781 | 29% | 1,781 | 86% |

Note. Scenario column headings refer to parameters on which clients and research participants must match. P = Problem; PAG = Problem-Age-Gender; PAGES = Problem-Age-Gender-Ethnicity-Setting. Ns refer to the total youths within the row label class; percentages refer to the percent of those youths not coverable.

Table 2

Treatments relevant to youths in the PAG scenario with coverage criterion of 100% of coverable youth, assuming the prior availability of a Multisystemic Therapy for disruptive behavior.

| Treatment Type | Practice or treatment in minimum set? | Applies to this % of sample youth | % of youths lost if practice or treatment dropped |
|--|---------------------------------------|-----------------------------------|---|
| Cognitive Behavior Therapy | ✓ | 43.4% | 8.9% |
| Self Verbalization | A | 8.8% | 1.8% |
| Social Skills and Medication | B | 7.0% | 1.8% |
| Contingency Management | ✓ | 16.5% | 1.3% |
| Parent Management Training | ✓ | 5.4% | 1.0% |
| Cognitive Behavior Therapy with Parents Included | ✓ | 27.7% | 0.3% |
| Intensive Communication Training | ✓ | 0.7% | 0.3% |
| Intensive Behavioral Treatment | ✓ | 0.6% | 0.2% |
| Parent Management Training and Problem Solving | ✓ | 7.2% | 0.1% |

Note. Two minimal sets were identified, and checkmarks indicate treatments found in both sets, whereas letters ‘A’ and ‘B’ indicate the two treatments of which only one is needed to complete a minimal set.

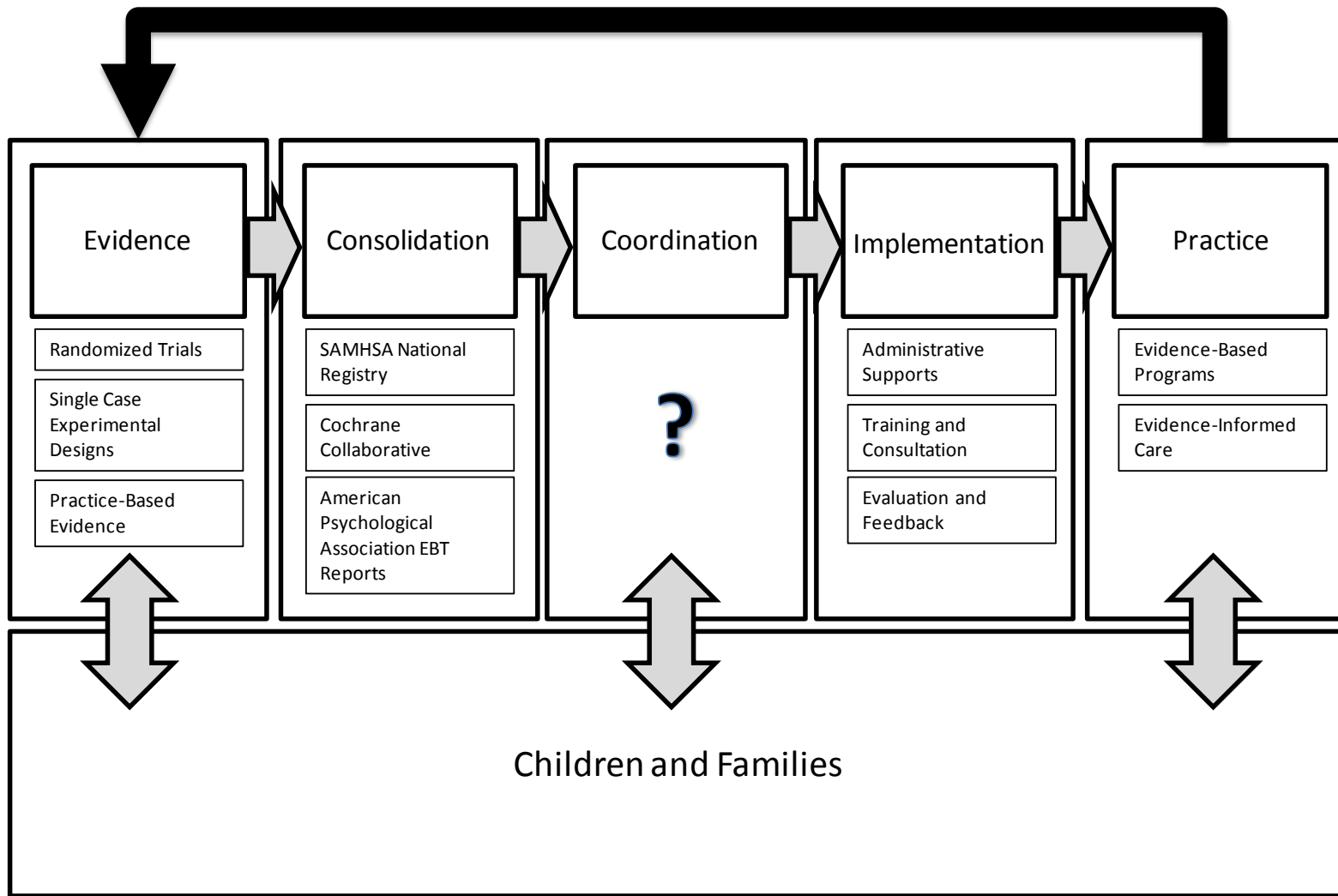
FIGURE CAPTIONS

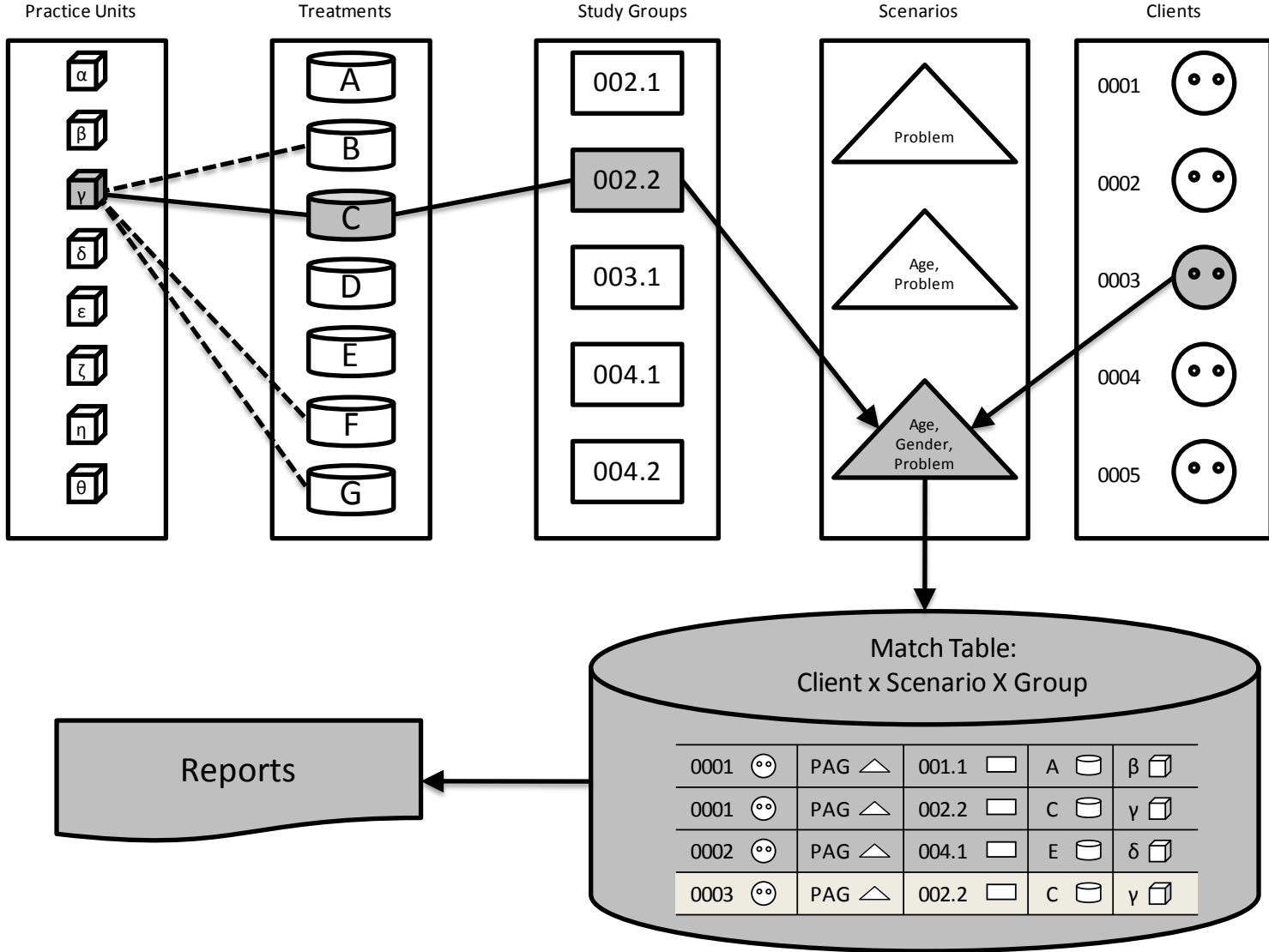
Figure 1. An illustration of the relation of evidence and practice in the context of children’s mental health services.

Note. Examples are listed for steps for which there are well-defined literatures, products, or concepts.

Figure 2. Graphical illustration of the relevance mapping algorithm for determining client coverage.

Note. Shaded elements in the top of the figure show the testing of a single client \times group \times scenario combination, recorded in the last row of the match table below. In the step of the analysis illustrated, client 3 matches a study group on problem, age, and gender, and that study group tested a specific manual used in the study (i.e., “treatment C”), which belongs to a class of treatments γ (for example, treatment C could be a specific CBT protocol, and γ could represent the class of all CBT protocols).





REFERENCES

- Agency for Health Research and Quality. (2001). *Translating research into practice (TRIP)-II*. Rockville, MD: Author. Available on the internet at <http://www.ahrq.gov/research/trip2fac.htm>
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.
- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*, 271-285.
- Barlow, D. H. (1981). On the relation of clinical research to clinical practice: Current issues, new directions. *Journal of Consulting and Clinical Psychology*, *49*(2), 147-155.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*, 7-18.
- Chorpita, B. F., Daleiden E. L., Ebesutani C. E., Young, J., Becker, K. D., et al. (in press). Evidence based treatments for children and adolescents: An updated review of efficacy and clinical utility. *Clinical Psychology: Science and Practice*.
- Collatz, L. & Wetterling, W. (1975). *Optimization Problems*. New York: Springer-Verlag, N. Y., Inc.
- Elliott, R. (Ed.). (1998). The empirically supported treatments controversy [special section]. *Psychotherapy Research*, *8*, 115-170.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

- Frenk, J. (2009, December). *Globalization and Health: The Role of Knowledge in an Interdependent World*. David E. Barmes Global Health Lecture. Lecture conducted from National Institutes of Health, Bethesda, MD.
- Glass, C. R., & Arnkoff, D. B. (1996). Psychotherapy integration and empirically validated treatments: Introduction to the special series. *Journal of Psychotherapy Integration*, 6(3), 183-189.
- Glisson, C., Gibbons, R., Green, E. P., Hoagwood, K., Jensen, P. S., Miranda, J., . . . Chorpita, B. F. (2008). Assessing the organizational social context (OSC) of mental health services: Implications for research and practice. *Administration and Policy in Mental Health and Mental Health Services Research*, 35(1-2), 98-113.
- Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J. M., Caswell, W., & Robinson, N. (2006). Lost in Knowledge Translation: Time for a Map? *Journal of Continuing Education in Health Professions*, 26(1), 13-24.
- Haynes R. B. (2001). Of studies, syntheses, synopses, and systems: The “4S” evolution of services for finding current best evidence. *ACP Journal Club*, 134, A11-A13.
- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic treatment for antisocial behavior in children and adolescents*. New York: Guilford Press.
- Hogan, M. F. (2003). The President's New Freedom Commission: Recommendations to transform mental health care in America. *Psychiatric Services*, 54, 1467-1474.
- Hromkovič, J. (2004). *Algorithmic adventures: From knowledge to magic*. New York: Springer.
- Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.

- Kazdin, A. E. (Ed.) (1996). Validated treatments: Multiple perspectives and issues [special section]. *Clinical Psychology: Science and Practice*, 3, 216-267.
- Kazdin, A. E. (Ed.) (1998). Treatment manuals in clinical practice [special section]. *Clinical Psychology: Science and Practice*, 5, 361-407.
- Kendall, P. C. (1999). Clinical significance [special section]. *Journal of Consulting and Clinical Psychology*, 67, 283-339.
- Kendall, P. C., Chambless, D. L. (Eds.) (1998). Empirically supported psychological therapies [special section]. *Journal of Consulting and Clinical Psychology*, 66, 3-167.
- Nathan, P. E., & Gorman, J. M. (Eds.). (1998). *A guide to treatments that work*. New York: Oxford University Press.
- Nathan, P. E., Stuart, S. P., & Dolan, S. L. (2000). Research on psychotherapy efficacy and effectiveness: Between Scylla and Charybdis? *Psychological Bulletin*, 126, 964-981.
- National Advisory Mental Health Council Workgroup on Services Research and Clinical Epidemiology (2006). *The road ahead: Research partnerships to transform services*. Washington, DC: Author.
- National Institute of Mental Health (2008). *National Institute of Mental Health strategic plan*. Bethesda: MD: Author.
- Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology. Special Issue: Evidence-Based Psychosocial Treatments for Children and Adolescents: A Ten Year Update*, 37(1), 8-38.
- Schuster, M. A., McGlynn, E. A., & Brook, R. H. (2005). How good is the quality of health care in the United States? 1998. *The Milbank Quarterly*, 83, 843-895.

- Shafran, R., Clark, D. M., Fairburn, C. G., Arntz, A., Barlow, D. H., Ehlers, A., . . . Wilson, G. T. (2009). Mind the gap: Improving the dissemination of CBT. *Behaviour Research and Therapy*, 47(11), 902-909.
- Silverman, W. K., & Hinshaw, S. P. (2008). The second special issue on evidence-based psychosocial treatments for children and adolescents: A 10-year update. *Journal of Clinical Child and Adolescent Psychology. Special Issue: Evidence-Based Psychosocial Treatments for Children and Adolescents: A Ten Year Update*, 37, 1-7.
- Society of Clinical Child and Adolescent Psychology (2011). *Evidence-based mental health treatment for children and adolescents*. Available on the internet at <http://www.abct.org/sccap/>
- Substance Abuse and Mental Health Services Administration (2006). *From exclusion to belonging: Transforming mental health care in America*. Rockville, MD: Author.
- Substance Abuse and Mental Health Services Administration (2011). *National registry of effective programs and practices*. Available on the internet at <http://nrepp.samhsa.gov/>
- Task Force on Promotion and Dissemination of Psychological Procedures, Division of Clinical Psychology, American Psychological Association. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, 48, 3–23.
- U.S. Department of Health and Human Services. (2007). *U.S. Department of Health and Human Services Strategic Plan*. Washington, DC: Author.
- VandenBos, G. R., (Ed.). (1996). Outcome assessment of psychotherapy [special issue]. *American Psychologist*, 51, 1005-1006.

Weisz, J., Hawley, K. M., & Jensen-Doss, A. (2004). Empirically tested psychotherapies for youth internalizing and externalizing problems and disorders. *Child and Adolescent Psychiatric Clinics of North America*, 13, 729-815.

Weisz, J. R., Hawley, K. M., Pilkonis, P. A., Woody, S. R., & Follette W. C. (2000). Stressing the (other) three Rs in the search for empirically supported treatments: Review procedures, research quality, relevance to practice and the public interest. *Clinical Psychology: Science and Practice*, 7, 243-258.

Zima, B.T., Hurlburt, M.S., Knapp, P., Ladd, H., Tang, L., Duan, N., . . . Wells, K. (2005). Quality of publicly-funded outpatient specialty mental health care for common childhood psychiatric disorders in California. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44, 130-144.

CHAPTER 2:

How We Define Treatment Affects Coverage and Efficiency in Coordinating Evidence Based

Practice

ABSTRACT

This study assessed the limits of coverage possible with evidence-based treatments (EBTs) as traditionally conceptualized even when all the programs listed in a large national report are assumed available in a clinical workforce. The study empirically examines how the way we define treatments affects major clinical questions such as the portion of youths who can be served by appropriate EBTs and the learning burden for service systems pursuing implementation. Treatments are conceptualized as (a) intact treatment *programs* (the current tradition in describing EBTs), and (b) collections of their constituent common procedures, referred to as *practice elements*. Programs listed by the National Registry of Effective Practices and Programs (NREPP) and the California Institute of Mental Health (CIMH) were selected and all available clinical trials testing the programs were analyzed. Practice elements were identified from these same studies and from studies of other treatments that met a standard of evidence but had not been organized into programs listed by NREPP or CIMH. Additionally, a hybrid model is introduced that uses combinations of programs and practice elements. Among a large, diverse clinical population, results identified 13% (for NREPP) and 19% (for CIMH) of children for whom practice elements provide an evidence-based treatment option when no EBT would otherwise be available. Further, the considerable efficiency advantage found for the practice element and hybrid models may be the difference between evidence based approaches reaching or not reaching an even larger number of treatment-seeking youths.

INTRODUCTION

Despite the formidable number of mental health treatments with sound empirical support, evidence-based treatments (EBTs) are still far from being consistently available to those in need. Continued treatment development and research to grow the evidence base, along with efforts to understand the factors that make implementation of EBTs successful, are vital to solving this problem. However, we have suggested that overcoming barriers in a related domain – service coordination – is equally essential and often overlooked (Chorpita, Bernstein, & Daleiden, in press). This paper focuses on the effects of a particular coordination concern that is subtle but may have dramatic clinical ramifications: how we define treatments. To realize the benefit of the numerous successfully tested treatments already identified, we need ways to leverage the entire evidence base. Yet service systems cannot afford to implement all available treatment programs, and the result is that most clients receive “usual care.” This study empirically examines how the way EBTs have traditionally been conceptualized – that is, the way we parse the evidence base into meaningful chunks – affects major clinical questions such as the portion of youths who can be served by appropriate EBTs, as well as the burden and feasibility of learning the most relevant EBTs for service systems pursuing implementation.

Approaches to Treatment Definition

The question of how best to define treatments has played an important role in the evidence-based practice movement of the past two decades, but has received relatively little explicit attention. Whereas early attempts to summarize therapy effects lumped treatments broadly together and thus failed to find outcome differences between them (e.g., Luborsky, Singer, & Luborsky, 1975; Smith, Glass, & Miller, 1980), subsequent reviews suggested that some differences exist between interventions when conceptualized at the level of their theoretical

background (e.g., advantages for behavior therapy for youths; Weisz, Weiss, Alicke, & Klotz, 1987). Efforts in the 1990s to systematically define the manner in which treatments are evaluated shifted the focus to specific manuals. A detailed definition of efficacy produced by the American Psychological Association (APA) Task Force on Promotion and Dissemination of Psychological Procedures (1995) provided criteria that a manual could satisfy to be labeled as empirical supported. Many researchers now design studies explicitly to meet these criteria, and an increasingly large number of interventions have met them—over 300 for children and adolescents (Chorpita et al., in press - b). These have been summarized in numerous lists of evidence-based treatments (EBTs; e.g., Substance Abuse and Mental Health Services Administration [SAMHSA], 2011; Silverman & Hinshaw, 2008). However, while the task force’s criteria defined interventions at the level of the manual, the summaries generally do not list specific manuals but rather group manuals into general classes based on global judgments about whether they share enough similarities to constitute a single generic treatment (Chorpita, Daleiden, & Weisz, 2005).

The practical implications of these different levels of analysis (e.g., specific manuals vs. general classes of treatment) have yet to be widely considered in research and policy initiatives. In discussions of evidence-based practice within mental health, the level of analysis used to describe treatment is often left implicit despite the lack of consensus noted above (e.g., National Advisory Mental Health Council [NAMHC] Workgroup on Services Research and Clinical Epidemiology, 2006; SAMHSA, 2006). One reason for the absence of attention may be that this definitional issue and its implications are not readily apparent or salient.

Levels of Analysis of Treatment Operations

As described above, there are currently multiple levels of specificity used to translate research knowledge about treatments into practice—and here we focus on three. First, we define treatment *programs* as descriptions of integrated treatments that are the products of specific research laboratories or investigators. Examples of treatment programs include the Coping Cat (Kendall, Kane, Howard, & Siqueland, 1990), Multisystemic Therapy (Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 1998), and the Incredible Years (Webster-Stratton & Reid, 2003). Treatment programs are typically the unit of analysis used to define treatments in large consolidated reviews—and they are often the implicit unit of analysis in traditional discussions of evidence-based practice within mental health (e.g., NAMHC Workgroup on Services Research and Clinical Epidemiology, 2006; SAMHSA, 2006).

A second level of analysis is treatment *protocols*, which we define as the manualized or structured sets of treatment instructions specific to a given study. Thus, within the Coping Cat program, there have been five different versions tested across five different randomized trials (Flannery-Schroeder & Kendall, 2000; Kendall, 1994; Kendall et al., 1997; Kendall, Hudson, Gosch, Flannery-Schroeder, & Suveg, 2008; Walkup et al., 2008). Because they are the closest representation of what was actually tested, treatment protocols are likely the most precise specification of the verified treatment procedures that are expected to lead to positive outcomes in future applications. Nevertheless, we know of no practice and policy recommendations that focus on specific protocols as the level of analysis. This is one indication that current traditions of treatment specification may be somewhat arbitrary or artifactual (e.g., consistent with treatment developers' frame of reference), rather than optimal.

A third level of analysis, whose merits (Chorpita, et al., 2005) and limitations (Chorpita, Becker, & Daleiden, 2007) we have discussed elsewhere, involves treatment *practices* (also

known as *practice elements*), which we describe as discrete procedures that are structured components of a larger course of treatment. Examples of practice elements include “Time Out,” “Relaxation Training,” and “Psychoeducation.” In a recent review of EBTs for children’s mental health problems (Chorpita & Daleiden, 2009), we coded over 600 treatment protocols for their component practices and showed that most evidence-based protocols within a particular problem area share a majority of practice elements. For example, most of the 21 evidence-based protocols for depression included the practice elements of cognitive restructuring, self-monitoring, pleasant activity scheduling, problem-solving training, and psychoeducation. The same pattern of practice element overlap was noted for treatment protocols for autism spectrum, anxiety, ADHD, disruptive behavior disorders, and substance use. These patterns, along with the work of numerous other investigators (e.g., Garland, Hawley, Brookman-Frazee, & Hurlburt, 2008; Rotheram-Borus et al., 2009; Ingram, Flannery, Elkavich, & Rotheram-Borus, 2008; Kaminski, Valle, Filene, & Boyle, 2008) have suggested an ability to aggregate and interpret research findings according to levels of analysis other than by treatment program.

When considering these different levels of analysis of treatment procedures, there clearly is a tradeoff between specificity – or the precision with which the procedures are represented – and efficiency – or the degree the knowledge about treatments can be organized into practical meaningful patterns. At either end of the spectrum, problems can occur. For example, with the exception of some computer-delivered protocols (e.g., Klingberg et al., 2005), most treatment protocols do not specify exactly what was said, who sat where, etc. This is presumably because at some level we know that such detail “overspecifies” the treatment (i.e., it includes more details than are believed to be essential for the desired outcome). At the other end of the spectrum, one can offer broad descriptions of the treatment approaches (e.g., “Cognitive Behavior Therapy”) or

summaries of the practice elements represented in the evidence base, which can be quite neatly consolidated, but may under-represent the detail needed for therapist to produce with fidelity the proper behaviors to lead to the desired outcomes. The specificity-efficiency tradeoff may have particular relevance to the larger issue of the application of evidence to practice precisely because the field has not yet examined the practical implications of these different levels of analysis.

Effects of Treatment Definition on Coverage and Coordination

The goal of the current study is to compare two of these approaches to defining treatment with regard to their effect on clinically important questions like how many clients can receive appropriate treatment and the efficiency of coordinating multiple treatments within an organization. We compare the program level of analysis, which represents the current tradition in national and state level initiatives, and the practice element level of analysis, in which treatment procedures are represented in terms of their individual component techniques. Because of their implicit nomination-based (as opposed to comprehensiveness-based) procedures for identifying treatments, and because many successful treatments were never organized as programs, national reports of EBTs (e.g., NREPP) and lists of EBTs supported by large state organizations (e.g., California Institute of Mental Health; CIMH) represent only a subset of all treatments in the literature that have empirical support. Thus, when analysis is limited to the programs included by these sources, we expect the portion of children in the service sample who are coverable by an EBT to decrease relative to when the literature is represented by its summary collection of practices coded from *all EBTs* in a comprehensive literature review. Further, we predict that conceptualizing treatments at the practice element level will offer substantially more efficient solutions to identifying the smallest number of treatments to serve the largest number of

children. The basis of this prediction about efficiency is that practice elements identify commonalities among many treatments in the literature.

In summary, this study aims to demonstrate clinical implications of defining treatment as intact programs versus component practices. Empirical analyses address the following three questions. First, how do the two approaches to treatment definition affect the portion of children in the service sample who are coverable by an EBT? Second, how efficient are the two approaches in terms of the number of treatments required to serve the coverable youths? Third, to what degree can practice elements cover children for whom no evidence-based programs are available on the consolidated lists? For the third aim, a hybrid model is introduced that uses combinations of programs and practice elements, more fully utilizing the literature so as to cover youths for whom no EBT would otherwise be available.

METHOD

Client Sample

The client data sample for this study comes from EMQ FamiliesFirst (EMQFF), a large mental health service agency with a service population well distributed between northern, southern and central California. This sample represents 3,793 youths receiving services from EMQFF between January 2009 and May 2010. For youths with multiple episodes of treatment with EMQFF, only the first episode during this time period was considered. The only inclusion criteria were data availability and age less than 19 years. To be included in analysis, a child had to have data on primary problem, age, and gender. The data availability criterion was used because these variables were required in all analyses considered in this report, and so otherwise the applicability of the research findings to a particular child could not be estimated. Of the 3,793 youths in the population, 2,186 were males (57.6%) and 1,607 were females (42.4%). Age of

youths in the sample ranged from 0 to 18 years with a median age of 13 (M=12.2, SD=4.1). Youth ethnic groups included Hispanic/Latino (n=1492; 39.3%), Caucasian (n=1238; 32.6%), African American (n=669; 17.6%), Asian American (n=189; 5.0%), Native American/Alaskan (n=31; 0.8%), and Multiethnic (n=16; 0.4%), with ethnicity data unavailable for 158 youths (4.2%).

Chart diagnoses were based on Diagnostic and Statistical Manual of Mental Disorders (4th ed. [DSM-IV]; APA, 2000) codes. For the purposes of this investigation, DSM-IV diagnoses present in the charts were cross-walked to 10 broad categories (the mapping of all diagnoses to problem groups is available upon request), shown in Table 1, and frequencies are reported under the column labeled "N." The decision to use this small number of broad problem categories was made to provide a common framework for matching clients to studies, given that research studies use a wide variety of taxonomies and methods for classifying and enrolling participants, not limited to diagnosis. Chorpita et al. (in press - a) describe the tradeoff involved in this decision between the precision with which a problem is characterized and the availability of data to inform a match. Given that less than half the studies in our review reported clinical diagnosis, we deemed the data availability concern to be more pressing and thus opted for the broad categories.

Study Datasets

In order to determine which treatments are candidates for analysis, relevance mapping requires the user to define or select a standard of evidence that identifies which treatments in the study dataset are considered evidence-based. For the comparisons central to the current study's aims, we have selected both a national example (NREPP) and a regional example (CIMH) of a consolidated list of EBTs to analyze in conjunction with a private-source database of randomized clinical trial (RCT) studies (PracticeWise Evidence Based Services database [PWEBS]);

PracticeWise, 2011). NREPP programs were identified by performing a search for mental health and substance abuse programs on the NREPP website (SAMHSA, 2011) and then limiting the results to those programs for which an RCT study could be identified using the process described below for mapping programs to studies (see *Procedure*). This method resulted in 24 treatment programs.¹ CIMH lists 8 supported treatment programs (CIMH, 2011), each of which was included in the analysis since each had been tested in at least one RCT.²

While the NREPP and CIMH lists provide criteria for determining which treatments to include in the analysis, additional criteria were needed to determine which studies of those treatments to consider as supporting evidence and thus to include in the study datasets. For example, consider a hypothetical program listed by NREPP that has been tested in 3 RCT studies. If two of the studies found the treatment to be effective with younger adolescents and the third tested the treatment with older adolescents and found it to be ineffective, it is important that the relevance mapping analyses only allow the treatment to cover younger adolescents. Since the NREPP and CIMH lists do not specify which studies to include (NREPP lists studies on which the NREPP review was based, but the list is not exhaustive or updated), an additional study standard is needed. We thus employed a study-level standard of evidence, based largely on the criteria developed and employed by the Task Force on Promotion and Dissemination of

¹ The 24 NREPP programs were: Adolescent Community Reinforcement Approach, Adolescent Coping With Depression, Brief Strategic Family Therapy, CARE (Care, Assess, Respond, Empower), Children's Summer Treatment Program, Cognitive Behavioral Intervention for Trauma in Schools, Cognitive Behavioral Therapy for Adolescent Depression, Coping Cat, Family Behavior Therapy, Family Support Network, Incredible Years, Multidimensional Family Therapy, Multidimensional Treatment Foster Care, Multisystemic Therapy for Juvenile Offenders, Multisystemic Therapy for Youth With Problem Sexual Behaviors, Multisystemic Therapy With Psychiatric Supports, Parent-Child Interaction Therapy, Parenting Through Change, Project ACHIEVE, Project ALERT, Seeking Safety, Teen Intervene, Trauma-Focused Cognitive Behavioral Therapy, and Triple P--Positive Parenting Program.

² The 8 CIMH programs were: Aggression Replacement Training, Depression Treatment Quality Improvement, Functional Family Therapy, Incredible Years, Multidimensional Treatment Foster Care, Multisystemic Therapy, Trauma Focused Cognitive Behavior Therapy, Triple P--Positive Parenting Program. Wraparound Community Development Team is also supported by CIMH but was not included in the analysis because it is not a specific manualized intervention, but rather is described as a "dissemination process" designed to support high fidelity implementation of other programs (CIMH, 2011).

Psychological Procedures (1995). According to this standard, a manualized treatment must (a) show statistically superior outcomes to a waitlist or no-treatment control group in at least two randomized trials, (b) show statistically superior outcomes to an active treatment or psychological placebo in at least one randomized trial, or (c) show equivalent outcomes to an already established evidence-based treatment in at least one randomized trial in which the average group size is at least 30 participants (see Chorpita et al., in press - b, for additional details). Table 1 summarizes the studies that met this standard from the full dataset as well as the subsets corresponding to NREPP and CIMH programs. Additional characteristics of a recent version of this dataset have been reported in previous research (Chorpita et al., in press - b).

Procedure

Linking programs and practice elements to studies. To determine which studies belonged in the study datasets for NREPP and CIMH, a mapping was created between treatment programs and their corresponding studies and protocols in the full PWEBS database. Because NREPP and CIMH each define their constituent programs somewhat differently (e.g., Multisystemic Therapy [MST] is considered to be a single program by CIMH, but NREPP lists MST as 3 different programs), mappings were created separately for each organization. As a first step in mapping between programs and studies, lists of study references for each program were obtained directly from the NREPP and CIMH websites where available, and otherwise from literature searches using electronic databases. The identified studies were then located in the PWEBS database and their protocols were recorded. A second round of mapping was performed in the reverse direction, with the first author examining all 939 protocols in the PWEBS database and recording mappings to programs from the NREPP and CIMH lists when applicable. This additional pass allowed for exhaustive identification of the protocols matching each program. Reliability of the

program coding was also examined. For each consolidated knowledge list, 20% of the manuals identified as mapping to programs, or a minimum of 10 manuals, were randomly selected to be mapped by an independent second expert rater (i.e., the second author). The resulting rates of agreement were 95% for NREPP ($\kappa = .94$) and 91% for CIMH ($\kappa = .88$).

For analyses using practice elements, protocols were included from the winning groups in all studies that met the standard of evidence described above. Two raters coded each of these protocols regarding the presence or absence of 59 practice elements, and an expert rater performed a final validation and review of all codes. A detailed description of the practice element coding and reliability is reported by Chorpita & Daleiden (2009).

Relevance mapping. The analytic procedures for this study followed the relevance mapping framework, described in detail elsewhere (Chorpita et al., in press - a). The central part of the relevance mapping model involves taking every client in a given service population and determining which published research trials have participants with matching characteristics. Each child in the client data sample is compared to each study in a study dataset to determine if the child is "covered" by that study. These structured comparisons can use any variables common to the client and study datasets, and in the current study, primary problem, age, and gender were used for matching children to studies (see *Data Analysis*). When a child is found to match a particular study on these parameters, a record of the child-study combination is added to a list of all matches. The resulting list of matches between children and studies is then summarized to answer questions about the characteristics of coverable and non-coverable youths.

Additionally, the list of matches is the starting point for a further set of optimization analyzes that aim to find the smallest sets of treatments that combine to cover the most youths. These best complimentary combinations are found by an automated search that tests many different

groupings of treatments to find the smallest sets that have matches in the list for a criterion percentage of youths. Details of the matching and the optimal treatment set identification analyses have been reported previously (Chorpita et al., in press - a).

Data Analysis

To address the first aim of the study, investigating how the two approaches to treatment definition (i.e., programs vs. practice elements) affect the portion of youths coverable by an EBT, coverability was analyzed using each of the three study datasets: (a) all studies for which practice elements had been coded (i.e., the full PWEBS dataset containing all identified randomized clinical trials), (b) only studies of programs cataloged in NREPP, and (c) only studies of programs supported by CIMH. For each of the three study datasets, only the studies that met the evidence criteria described above were entered into the analysis. Relevance mapping analyses were performed under a "problem-age-gender" scenario, defining an EBT as relevant to clients with the same primary problem, within the same age range, and with the same gender as the participants in a study in which that treatment was successfully tested.³ These parameters were selected to represent what may be common assumptions about the considerations most important for treatment generalizability, since little empirical information is available to inform what parameters may be essential for generalizability (Chorpita et al., in press - a).

Characteristics of non-coverable youths were examined by counting the number of non-coverable youths with each primary problem area.

For the study's second aim, evaluating the relative efficiency of approaching treatment as programs versus practices, minimization analyses were performed as described above (see

³ For the "problem" parameter, coverage required the problem area targeted for treatment in a given study and used for participant inclusion in that study to match the primary problem area of a child in the client dataset. For example, a client with a primary problem area of depression could only be covered by treatments from studies that targeted the treatment of depression and included youth with depression as a primary problem. For age and gender, to cover a given client a study was required to have at least one participant with matching values.

relevance mapping) for practice elements, NREPP programs, and CIMH programs. Minimum sets of programs and practices were then compared with regard to their associated learning burden. The number of practice elements contained in a minimum set was used as a proxy for the learning burden placed on a system pursuing implementation. To provide a common metric for comparison between the minimum sets of practice elements and programs, learning burden for programs was calculated in two ways: (a) by summing the number of practice elements corresponding to each program in a minimum set, and (b) by counting the unique practice elements corresponding to the programs in a minimum set (see *Discussion* for the assumptions underlying this approach). Since many programs corresponded to multiple protocols (as described in the introduction and summarized in Table 1A), for summing, the mean number of practice elements was used from among the protocols that met the evidence criteria and thus entered into the analysis. The aim of the learning burden comparisons is to provide a comparative analysis of the difficulty (and to some degree feasibility) of actually achieving the optimized coverage identified under these different assumptions about how best to parse the evidence base.

The study's third aim is to evaluate the prospect of using practice elements to cover children for whom no evidence-based programs are available on the consolidated lists. To address this aim, the client dataset was "residualized" by removing youths covered by any program on a consolidated list. Separate residualized client datasets were created for NREPP and CIMH. These residualized client datasets were then subjected to coverability analysis and identification of minimum sets of practice elements, just as described above for the full client dataset.

RESULTS

Chorpita et al. (in press - a) describe two main categories of relevance mapping reports that can be produced once the dataset of client to study group matches is developed. The first category is *client reports* which refer to clients as dependent variables (e.g., % of clients coverable) and expose who in a given service population is or is not coverable. Client reports answer the question, for whom in the service sample is there *any* relevant EBT? The second category is *practice reports* which refer to practice units as dependent variables and answer questions such as, how many and which practice units are required to cover as many clients as possible?

Aim 1: Differences in coverage

Table 2 presents a client report examining coverage in the EMQFF dataset when youths were required to match research participants on primary problem, age, and gender. The first row shows that less clients were left uncovered when analyzed in the context of the full literature from which practice elements were coded (37.1% not covered) versus when the study dataset included only studies of NREPP programs (49.9% not covered) or CIMH programs (56.2% not covered). As can be seen in the problem type rows of the table, depression was the most common problem area (n = 1,055); as an example, let us examine this area more closely. Only 2% of youths with a primary problem of depression were not covered by practice elements, whereas 16% and 22% were not covered by NREPP and CIMH programs, respectively. Thus, though two depression-focused programs were available on the NREPP list (Adolescent Coping With Depression [CWD-A] and CBT for Adolescent Depression) and one was available on the CIMH list (Depression Treatment Quality Improvement; DTQI), these were not sufficient to provide coverage for a considerable number of the youths in need of treatment for depression. The finding that the full study dataset represented by practice elements covers many more depressed

youths than these programs may not be surprising in light of the information in Table 1, which indicates that the full study dataset contained 23 RCTs that met the standard of evidence for depression while the NREPP and CIMH programs corresponded to only 5 and 1 RCTs, respectively. The main effect of age within the depression problem area is a more specific explanatory factor: whereas the 23 depression studies in the practice element dataset combined to cover ages 8 to 18, none of the NREPP depression studies covered youths below age 12, and the CIMH study did not cover youths below age 13. Since 140 depressed youths were ages 8 to 11, this difference explains most of the coverage loss. Additional lost coverage may be due to the *combination* of age and gender criteria within depression studies, since the same study was required to match a child on both of these criteria, as well as problem area. Along with this difference in coverage of youths with depression, similar differences are notable for the problem areas of anxiety, attention, autism spectrum, and traumatic stress.

The largest category of uncovered youths was “other problems.” 1,027 youths were in this category and thus could not be covered by any of the models. The most common diagnoses for youths in this category were: adjustment disorders (n = 303),⁴ other mood disorders (n = 259), anxiety disorder not otherwise specified (NOS; n = 156),⁵ disorders of infancy or early childhood NOS (n = 118), other psychotic disorders (n = 64), intermittent explosive disorder (n = 51), and reactive attachment disorder (n = 39). Also, by definition the 63 youths whose primary diagnosis (and resulting primary problem) was “none” could not be covered by any treatment, since matching on primary problem was required. Finally, the columns labeled “NREPP+PE” and “CIMH+PE” are described under aim 3, below.

⁴ As with all diagnoses, adjustment disorders were mapped to specific problem areas when the categorization could be made unambiguously (e.g., adjustment disorder with anxiety, adjustment disorder with depressed mood) and were mapped to “other problems” when ambiguity prevented a single mapping (e.g., adjustment disorder with mixed anxiety and depressed mood, adjustment disorder unspecified).

⁵ The diagnosis anxiety disorder NOS is ambiguous in that it may represent either anxiety or trauma.

Aim 2: Efficiency of programs and practice elements

Tables 3 and 4 present practice reports for programs and practice elements, respectively. Table 3 shows that it would be possible to serve the full 50% of youths coverable by any NREPP program using just 8 of those programs, while 6 CIMH programs are needed to achieve the maximum 44% coverage possible with that list. The first two program rows show that CWD-A and DTQI alone are each applicable to more than 20% of clients, accounting for nearly half of the clients to whom any program in their respective minimum sets applied. The rightmost column for each list refers to the percentage of youths who would no longer be coverable were a particular approach to be dropped from the identified service array. The table shows that there is very little overlap in coverage among the programs within the identified arrays, since in most cases the percentage of youths who would be lost if a program were dropped is the same as the percentage of youths to whom that program is applicable. Other treatments listed in the table thus provide *complementary* coverage, in many cases with the identified programs corresponding to a particular problem area (e.g., Coping Cat for anxiety) or problem area and age combination (e.g., Incredible Years for disruptive behavior in younger youths) addressed by none of the other programs in the array.

The first column of Table 3 shows the mean number of practice elements corresponding to each program, and the last rows total these values for each list in two different ways so as to provide a measure of total learning burden for the identified service array. The first total row simply sums the number of practice elements from each program in the array. The total learning burden for the NREPP programs using this metric ranges from the equivalent of 89.8 to 107.3 practice elements (depending on which of the minimal options is chosen), while the total burden for the CIMH programs is 82.1 practice elements. The second total row provides a count of the

unique practice elements among the programs in each array. Using this metric, the total for NREPP is 35 to 42, while the total for CIMH is 32 to 42. The first column grouping of Table 4 provides the same information for the practice element minimization analysis. 24 practice elements were sufficient to serve the full 63% of coverable youths. Correcting for the differences in coverage to allow for a more direct burden comparison, the practice element model requires 1.0 practice elements per 100 cases covered, whereas the CIMH models require the equivalent of 4.9 practice elements per 100 cases covered, and the NREPP model ranges from 4.6 to 5.4 practice elements per 100 cases covered.

Continuing to examine the first column grouping of Table 4, the practice element rows show that there is considerable redundancy within the minimal set, in that all practice elements apply to a much higher percentage of youths than would be lost to coverage if the practice element were dropped from the service array. Cognitive and child psychoeducation were the practices that uniquely applied to the most clients in the sample. If these practice elements were removed from the array, 12% and 9% of the youths would be lost to coverage, respectively, whereas in total each of the two practices applies to 55% of the sample (compared to 63% to which any practice element applies).

Aim 3: A hybrid model

The rightmost two columns of Table 2 and the rightmost two column groupings of Table 4 show results for the hybrid models that combine programs and practice elements. Table 2 shows that for both the NREPP and CIMH hybrid models, coverage is identical to the coverage provided by practice elements alone. This finding makes sense in light of the fact that the clients covered by NREPP and CIMH programs are a subset of those covered by practice elements. The

addition of practice elements thus newly covers the remainder of clients who were not coverable by programs but were by practice elements.

In Table 4, the hybrid columns show that the practice elements found to cover the additional clients are a similar but slightly reduced set as compared to those identified for the full sample. Twenty-one and 22 practice elements were needed to cover the 13% and 19% of the sample uniquely coverable by practice elements for NREPP and CIMH, respectively. Several different minimal set options were found both for the NREPP and CIMH hybrid models. Notably, all the “required” practice elements – those that appeared in each minimal set for a given list (i.e., the check-marked practice elements) – were identified in the practice elements-only model as well, indicating consistency in the practice elements found to most efficiently serve different subsets of the EMQFF client sample. Regarding learning burden, the table’s bottom rows combine the hybrid models’ learning burden from programs (detailed in Table 3) with the count of additional practice elements. The combined count for the NREPP hybrid model ranges from 114.2 to 135.2 (4.8 to 5.7 practice elements per 100 cases covered), and the total for the CIMH hybrid model is 107.9 (4.5 practice elements per 100 cases covered).

DISCUSSION

This study assessed the current limits of coverage afforded even when all the programs listed in a large national report are assumed to be available in a clinical workforce. The results show an approach to covering many of the children for whom no evidence-based treatment would otherwise be available. A primary result of this paper may thus be to make salient the considerable number of youths who would receive usual care in a system where only programs are available but who could otherwise be covered by common elements of EBTs. 483 children (13%) fell into this category when assuming the availability of all programs listed by NREPP,

and 722 children (19%) when the available programs were those supported by CIMH for the state of California. In real world service systems, where resource constraints make it impossible to implement all programs on these lists (or even the optimal set of programs identified by a relevance mapping analysis), the number of children that fall into this category may be much higher.

The other primary result of this paper is the finding of major differences in efficiency between the practice element and program models. Comparing burden using the metric examined in the Results section, practice elements per case covered, the practice element model was a multiple of 5 to 6 times more efficient than the program models.

Regarding the hybrid models, the findings indicate that the primary gain from the hybrid approach – as modeled for the current analysis – is increased *coverage* rather than efficiency. Only 2 to 3 less practice elements were needed to cover the “last” 13% and 17% of cases not already coverable by programs than were found necessary to cover the full 63% of coverable youths in the practice-elements-only model. Nonetheless, the hybrid models were slightly more efficient than their program-only alternatives. One sensible adaptation in light of these findings may be to complement the identified set of programs with all 24 practice elements identified for the practice-elements-only model, thereby providing a second treatment option (i.e., practice elements) for all the youths coverable by programs with little additional burden.

Another hybrid option is to maintain maximal coverage but improve *efficiency* rather than redundancy by decreasing the number of programs in the service array. The current analysis required 100% of program-coverable clients to be covered by programs. Approaches aimed at covering a somewhat smaller percentage of clients with programs (and thus uniquely covering more clients with practice elements) are likely to yield dramatically lower burden. For example,

it can be seen from Table 3 that the last two programs for both NREPP and CIMH uniquely cover less than 1% of clients. Dropping these programs from the service array and covering the small number of additional clients using practice elements would considerably increase efficiency of a hybrid solution.

In discussing the marked differences in efficiency found between the practice element and program based models, it is important to understand the assumptions underlying each model. Implicit to the program approach is an *assumption that each program is unique*. Training practices reflect this assumption in that a provider being trained in a given program is required to newly learn all aspects of that program. Thus a therapist who previously has been trained in the time out procedure from one disruptive behavior program, such as Incredible Years, does not “skip out” of learning the time out procedure when receiving training for a different disruptive behavior program, such as Triple P. Rather, the therapist receives training in the second program’s time out procedure with the same degree of detail as his or her fellow trainees who may not have been previously exposed to time out. This “blank slate” approach decreases logistical complexity for a training organization (e.g., it is easier to have all trainees attend the same two day training), but increases burden on providers with regard to time spent training and effort learning and remembering numerous unique programs. The assumption of uniqueness is in contrast to an *assumption of efficiency* utilized by the practice element model. The common elements approach assumes an acceptable degree of commonality between corresponding elements of different protocols. To be sure, real differences exist between the elements as implemented in specific protocols or programs. However, by opting to focus on what is shared in common, the practice elements approach removes the need to newly learn a clinical strategy for each treatment in which it appears.

In reality, the assumptions of uniqueness and efficiency likely represent extremes on a continuum. Indeed, some degree of additional learning burden is required to use a familiar practice element in a new context. And similarly, when being trained in a second program there is some degree of efficiency in re-learning a procedure like time out, even though the training generally proceeds as if “from scratch.” The results in Table 3 showed that the practice element model was considerably more efficient than the program models even when the efficiency assumption was applied to both (that is, when both were assumed to be at the “efficiency” end of the continuum). While the true position of these models on the continuum from efficiency to uniqueness is unknown, the current results give a picture of the possible range, with programs showing to require between a third and several fold more learning burden than practice elements.

The difference in efficiency may be of great importance since the learning burden that this study’s results show to be required of an organization attempting to use only programs while serving clients with treatments that match their needs may be beyond what is feasible for most real world service systems. Regardless of theoretical coverage maxima, the efficiency gained by a practice element or hybrid model may be the difference between evidence-based approaches reaching or not reaching an even larger number of treatment-seeking youths.

It is essential to emphasize that there are many pros and cons to approaching treatment as programs versus practice elements apart from those explored in this paper. Other issues are also highly relevant to a service system coordinating its treatment array, but beyond the scope of this paper, such as effectiveness, appeal to therapists, expiration and updateability of treatment procedures, compatibility with existing organizational infrastructure, and availability of systems and organizations for dissemination and training. Packaged programs and practice elements each have relative strengths and weaknesses in these areas, many of which have been examined in

detail elsewhere (e.g., Chorpita, et al., 2005; Chorpita, et al. 2007). This paper showed that the two approaches are not mutually exclusive but can be used in a complementary fashion like the hybrid models analyzed here.

Notably, the identity of the specific programs and practice elements identified as optimizing coverage in Tables 3 and 4 is not particularly important. Relevance mapping is inherently “local” in its approach, in that the solutions returned are a function of the mix of client characteristics inputted via the client dataset. The treatments listed thus represent what is most relevant to EMQFF, based on the clients they served in 2009 - 2010. For example, depression practices were found to be the most broadly applicable because depression was the most common problem area in the sample. The message of greater general import, in addition to the study's main findings comparing different ways to parse the evidence base, is the demonstration that from a vast number of possible combinations, the relevance mapping methodology can precisely identify treatment combinations that are optimal for a local population based on flexible assumptions regarding what is important in matching children to treatments.

A salient result not yet mentioned is the finding that more than 30% of children were covered by nothing. For more than a third of the youths in this client dataset – which represents a substantial portion of the nation’s largest state – no treatment from among the more than 500 RCTs reviewed was a match. This finding replicates the results of Chorpita et al. (in press - a), who used an unrelated client dataset from the Hawaii public mental health system, containing youths with a considerably different mix of problems. While some of the coverage failure is an artifact of client diagnoses that cannot be unambiguously mapped to the broad problem categories used in these studies, it is clear that the evidence base simply does not yet have answers for many common problems (cf. Schiffman, Becker, & Daleiden, 2006).

However, it is also important to note that the portion of youths found uncoverable is a function of the starting assumptions of the relevance mapping analyses. One set of assumptions involved our requirement that children match research participants on primary problem, age, and gender. Relaxing these requirements, for example, by removing age from the list of matching parameters, would considerably increase the number of youths coverable. While consideration of multiple scenarios representing different assumptions is beyond the scope of this paper, it would certainly be recommendable for the coordination process of a practice organization. Further, consideration of multiple scenarios provides important clues for how to proceed in treatment for the “uncoverable” youths. Chorpita et al. (in press - a) describe various options for providing evidence informed treatment for youths in this category. For example, treatments found to be generalizable when the age requirement is relaxed could be adjusted to be age appropriate for youths who had no matching treatment in this paper’s analysis (i.e., adaptation) or they could be used unmodified with this new population (i.e., extension). These options apply both to practice elements and programs.

A related consideration is that we used relatively liberal assumptions in matching youths to studies (e.g., a small number of broad problem categories, and matching on gender when a study contained at least one participant of a client's gender). Replacing these with more conservative assumptions (e.g., matching on precise diagnosis, and requiring a client's gender to match a minimum of 30% of a study's participants) would certainly decrease the overall number of youths coverable, and would likely also have the effect of broadening the coverage and efficiency gaps between practice elements and programs, since a proportional reduction in relevant literature would tend to have a larger effect on evidence lists that are smaller to begin with.

Other limitations of our analyses include the manner in which optimal treatment sets and burden were operationalized. The programs and practices reported in Tables 3 and 4 were identified by the analyses because they were the smallest groupings that covered the largest number of children. However, minimal size is just one component of optimality. Chorpita et al. (in press - a) describes the challenge of estimating “maximal expected impact” on a service population, and weighs alternative definitions such as predicted effect sizes. Regarding learning burden, we used the mean number of practice elements mapped to a program’s protocols as a proxy for burden, and the earlier portion of this Discussion section details the assumptions behind that choice. However, though the number of practice elements involved in a program appears to be a good indicator of burden for the therapists who must learn those programs, it is not a meaningful measure of other types of system burden, such as administrative burden, for which compatibility and complexity of chosen treatments may be more essential. Indeed, burden minimization requires deciding what or for whom you are trying to optimize and different local preferences may yield different results.

In summary, we found that even when all treatments from a large national list were assumed available, there was a sizable group of youths who were not coverable by programs but were coverable by the same evidence base, approached in a somewhat different configuration, i.e., practice elements. Moreover, the learning burden required to serve the coverable clients with appropriate treatments was strikingly lower for practice elements. Interestingly, our main finding regarding coverage might have been readily inferable even without the sophisticated matching analyses used in this study. That is, working with a subset of the evidence restricts the portion of youths who can be served by EBTs, and even a large national registry like NREPP corresponds to just a small subset of the full literature. What relevance mapping adds to this story is

information about the scope of that loss. In that regard, the results for CIMH similarly contribute by demonstrating that the extent to which a list truncates the available evidence is commensurate to size of the corresponding effect on coverage. The rapid continued growth of the evidence base and evidence-based programs present wonderful opportunities to alleviate suffering from mental illness, and also major challenges for service coordination. To realize the benefit of this abundance, we need ways to leverage the entire evidence base while keeping burden to a manageable level. The current study's results show that common elements and hybrid program–practice element approaches are two viable means of achieving this goal.

Table 1.

Study dataset characteristics. All studies in the study datasets were randomized clinical trials (RCTs). Protocols represent the actual manuals or other descriptions of treatment tested in an RCT, while a study group represents those participants in a study who received a specific treatment protocol.

A. Summary

| | Total count for category (count meeting evidence criteria) | | |
|--------------------------|--|----------|---------|
| | All | NREPP | CIMH |
| RCT studies ^a | 524 (255) | 96 (52) | 62 (34) |
| Protocols | 939 (288) | 103 (49) | 67 (31) |
| Study groups | 1295 (318) | 125 (63) | 82 (41) |

B. RCT studies by problem type

| Problem type | Total count of RCTs (count of RCTs meeting evidence criteria) ^a | | |
|---------------------|--|---------|---------|
| | All | NREPP | CIMH |
| Anxiety | 106 (75) | 5 (3) | 0 (0) |
| Attention | 58 (24) | 4 (2) | 2 (1) |
| Autism spectrum | 30 (10) | 1 (1) | 1 (1) |
| Depression | 37 (23) | 5 (5) | 1 (1) |
| Disruptive behavior | 145 (89) | 40 (28) | 35 (26) |
| Eating | 11 (4) | 0 (0) | 0 (0) |
| Mania | 3 (1) | 0 (0) | 0 (0) |
| Substance use | 31 (19) | 13 (8) | 4 (1) |
| Traumatic stress | 20 (9) | 8 (4) | 4 (3) |

^a RCT studies are listed as meeting evidence criteria if they tested at least one protocol that met the evidence criteria.

Table 2.

Children not coverable by evidence-based treatments identified in published randomized clinical trials (RCTs) corresponding to programs listed by the National Registry of Effective Practices and Programs (NREPP), the California Institute of Mental Health (CIMH), and by common practice elements (PEs) coded from all identified RCTs. Coverage required clients to match research participants on primary problem, age, and gender.

| | N | % Not Covered | | | | |
|---------------------|-------|---------------|-------|-------|----------|---------|
| | | PE | NREPP | CIMH | NREPP+PE | CIMH+PE |
| Total | 3,793 | 37.1% | 49.9% | 56.2% | 37.1% | 37.1% |
| Problem type | | | | | | |
| Depression | 1,055 | 2% | 16% | 22% | 2% | 2% |
| Disruptive behavior | 657 | 0% | 6% | 6% | 0% | 0% |
| Attention | 391 | 34% | 73% | 100% | 34% | 34% |
| Traumatic stress | 275 | 0% | 26% | 26% | 0% | 0% |
| Mania | 148 | 89% | 100% | 100% | 89% | 89% |
| Anxiety | 103 | 2% | 39% | 100% | 2% | 2% |
| Autism spectrum | 67 | 40% | 76% | 76% | 40% | 40% |
| Substance use | 5 | 0% | 0% | 40% | 0% | 0% |
| Eating | 2 | 0% | 100% | 100% | 0% | 0% |
| None | 63 | 100% | 100% | 100% | 100% | 100% |
| Other problems | 1,027 | 100% | 100% | 100% | 100% | 100% |

Table 3.

Programs relevant to youths with coverage criterion of 100% of coverable youths.

| Cases coverable | NREPP Programs | | | | CIMH Programs | | |
|---|----------------|-------------------------|-----------------------------------|-------------------------------------|-------------------------|-----------------------------------|-------------------------------------|
| | 1901 (50%) | | | | 1662 (44%) | | |
| Minimum programs | 8 | | | | 6 | | |
| Program | Number of PEs | Program in minimum set? | Applies to this % of sample youth | % of youths lost if program dropped | Program in minimum set? | Applies to this % of sample youth | % of youths lost if program dropped |
| CWD-A | 12.3 | ✓ | 23.5% | 23.5% | | | |
| DTQI | 18.0 | | | | ✓ | 21.7% | 21.7% |
| MST-JV, MST-PSB, or MST-Psychiatric | 17.0 to 23.5 | ✓ | 12.3% | 12.3% | | | |
| TF-CBT | 8.0 | ✓ | 5.4% | 5.4% | ✓ | 5.4% | 5.4% |
| MST | 18.6 | | | | ✓ | 12.3% | 5.2% |
| STP | 16.0 | ✓ | 2.8% | 2.8% | | | |
| Incredible Years | 12.0 | ✓ | 3.9% | 2.3% | ✓ | 3.9% | 2.3% |
| Coping Cat | 11.5 | ✓ | 1.7% | 1.7% | | | |
| Triple P | 11.0 | ✓ | 2.0% | 0.5% | ✓ | 2.0% | 0.5% |
| A-CRA, Brief Strategic Family Therapy, Family | 2.0 to 13.0 | ✓ | 0.1% | 0.1% | | | |

| | | | | |
|------------------------------|------|---------------|---|--------------|
| Behavior Therapy, or MDFT | | | | |
| Functional Family Therapy | 14.5 | | ✓ | 7.2% 0.1% |
| <hr/> | | | | |
| Total learning burden: | | | | |
| Sum of PEs | | 89.8 to 107.3 | | 82.1 |
| Unique PEs | | 35 to 42 | | 32 to 42 |
| <hr/> | | | | |

Note. Multiple minimal sets were identified for NREPP programs. For the rows listing groupings of programs, any one program from the grouping can be selected to complete a minimal set. PE = practice element; A-CRA = Adolescent Community Reinforcement Approach; CWD-A = Adolescent Coping With Depression; DTQI = Depression Treatment Quality Improvement; MDFT = Multidimensional Family Therapy; MST = Multisystemic Therapy (JV = for Juvenile Offenders; PSB = for Youth With Problem Sexual Behaviors; Psychiatric = With Psychiatric Supports); STP = Children's Summer Treatment Program; TF-CBT = Trauma-Focused Cognitive Behavioral Therapy.

Table 4.

Practice Elements (PEs) relevant to youths with coverage criterion of 100% of coverable youths.

| | PEs | | | NREPP programs and PEs (hybrid) | | | CIMH programs and PEs (hybrid) | | |
|---|----------------|-----------------------------------|--------------------------------|---------------------------------|-----------------------------------|--------------------------------|--------------------------------|-----------------------------------|--------------------------------|
| Cases coverable (all) | 2384 (63%) | | | 2384 (63%) | | | 2384 (63%) | | |
| Coverable by programs | | | | 1901 (50%) | | | 1662 (44%) | | |
| Coverable by PEs only | | | | 483 (13%) | | | 722 (19%) | | |
| Minimum PEs | 24 | | | 21 | | | 22 | | |
| Minimum programs | | | | 8 | | | 6 | | |
| PE | PE in min set? | Applies to this % of sample youth | % of youths lost if PE dropped | PE in min set? | Applies to this % of sample youth | % of youths lost if PE dropped | PE in min set? | Applies to this % of sample youth | % of youths lost if PE dropped |
| Cognitive | ✓ | 54.9% | 12.1% | ✓ | 8.0% | 3.9% | ✓ | 11.5% | 4.9% |
| Psychoeducation-Child | ✓ | 55.0% | 9.3% | ✓ | 8.4% | 3.6% | ✓ | 11.9% | 3.8% |
| Exposure | ✓ | 19.2% | 7.9% | ✓ | 2.8% | 2.4% | ✓ | 4.5% | 2.8% |
| Problem Solving | ✓ | 60.4% | 5.1% | ✓ | 12.1% | 3.4% | ✓ | 18.4% | 4.6% |
| Maintenance/Relapse Prevention | ✓ | 59.1% | 3.1% | ✓ | 9.4% | 1.6% | ✓ | 15.7% | 1.8% |
| Psychoeducation-Parent | ✓ | 62.1% | 3.0% | ✓ | 12.1% | 2.8% | ✓ | 18.4% | 3.6% |
| Social Skills Training | ✓ | 54.1% | 2.6% | ✓ | 9.4% | 2.2% | ✓ | 15.7% | 2.4% |
| Stimulus Control or Antecedent Management | ✓ | 52.6% | 1.9% | E | 6.2% | 0.9% | A,F,K,N | 10.7% | 1.0% |

| | | | | | | | | | |
|-----------------------------------|---|-------|------|-------|-------|------|---------------------------------|-------|------|
| Activity Scheduling | ✓ | 35.6% | 1.7% | ✓ | 6.4% | 2.9% | ✓ | 9.3% | 3.3% |
| Communication Skills | ✓ | 54.9% | 1.7% | ✓ | 7.9% | 1.6% | ✓ | 11.5% | 1.8% |
| Goal Setting | ✓ | 53.0% | 1.7% | | | | L,M,N | 9.6% | 0.1% |
| Parent Coping | ✓ | 47.7% | 1.6% | ✓ | 5.9% | 1.6% | ✓ | 9.4% | 1.7% |
| Assertiveness Training | ✓ | 50.9% | 1.5% | | | | A | 11.1% | 0.1% |
| Peer Pairing | ✓ | 34.2% | 1.4% | ✓ | 3.8% | 1.3% | ✓ | 7.9% | 1.8% |
| Relaxation | ✓ | 56.5% | 1.4% | A,C | 9.8% | 0.1% | B,D,G, I | 15.5% | 0.1% |
| Tangible Rewards | ✓ | 58.6% | 1.4% | ✓ | 11.4% | 1.1% | ✓ | 17.7% | 1.7% |
| Self-Monitoring | ✓ | 51.3% | 1.3% | B,D,E | 9.1% | 0.4% | A,C,E, F,H,J, K,L,M, N | 14.2% | 0.5% |
| Modeling | ✓ | 61.2% | 1.2% | ✓ | 11.5% | 1.2% | ✓ | 17.8% | 1.8% |
| Monitoring | ✓ | 30.9% | 1.0% | ✓ | 5.2% | 0.7% | ✓ | 9.2% | 1.3% |
| Praise | ✓ | 33.2% | 1.0% | ✓ | 7.1% | 1.1% | ✓ | 11.5% | 1.7% |
| Relationship/ Rapport Building | ✓ | 55.9% | 1.0% | ✓ | 7.6% | 0.7% | ✓ | 13.9% | 1.3% |
| Response Cost | ✓ | 23.9% | 1.0% | ✓ | 4.8% | 1.1% | ✓ | 7.6% | 1.7% |
| Insight Building | | | | A,B | 6.3% | 0.9% | B,C,G, H,L | 11.1% | 0.9% |
| Self-Reward/Self- Praise | | | | C,D | 7.2% | 0.9% | D,E,I,J ,M | 13.4% | 0.9% |
| Biofeedback/ Neurofeedback | ✓ | 4.9% | 0.3% | ✓ | 2.4% | 0.3% | ✓ | 4.9% | 0.3% |

| | | | | | | | | | |
|--|---|-------|------|---|------|----------------|-----------------------------------|-------|------|
| Family Engagement | | | | | | | B,C,D, E,F | 1.1% | 0.1% |
| Family Therapy | A | 42.9% | 0.1% | ✓ | 3.6% | 0.1% | A,B,C, D,E,F, G,H,I,J ,K | 6.8% | 0.1% |
| Motivational Interviewing | B | 17.4% | 0.1% | | | | G,H,I,J ,K,L,M ,N | 0.1% | 0.1% |
| Learning burden from PEs | | 24 | | | | 21 | | 22 | |
| Learning burden from programs | | 0 | | | | 89.8 to 107.3 | | 82.1 | |
| Total learning burden from PEs and programs | | 24 | | | | 110.8 to 128.3 | | 104.1 | |

Note. Checkmarks indicate practice elements found in all minimal sets for the column whereas letter (‘A’ through ‘N’) indicate practice elements found only in some minimal sets. A complete minimal set requires all check-marked practice elements in a column plus all practice elements with any one letter (e.g., all practice elements with the letter ‘A’ in that column).

REFERENCES

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.
- California Institute of Mental Health (2011). *Values-driven evidence-based practice in child and family services*. Available on the internet at <http://www.cimh.org/Services/Child-Family/Evidence-Based-Practice.aspx>
- Chorpita, B. F., Becker, K. D., & Daleiden, E. L. (2007). Understanding the common elements of evidence-based practice: Misconceptions and clinical examples. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(5), 647-652.
- Chorpita, B. F., Bernstein, A., & Daleiden, E. L. (in press - a). *Empirically Guided Coordination of Multiple Evidence-Based Treatments: An Illustration of Relevance Mapping in Children's Mental Health Services*. *Journal of Consulting and Clinical Psychology*.
- Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology, 77*(3), 566-579.
- Chorpita, B. F., Daleiden E. L., Ebesutani C. E., Young, J., Becker, K. D., et al. (in press - b). Evidence based treatments for children and adolescents: An updated review of efficacy and clinical utility. *Clinical Psychology: Science and Practice*.
- Chorpita, B. F., Daleiden, E., & Weisz, J. R. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research, 7*, 5-20.

- Flannery-Schroeder, E. C., & Kendall, P. C. (2000). Group and individual cognitive-behavioral treatments for youth with anxiety disorders: A randomized clinical trial. *Cognitive Therapy and Research, 24*(3), 251-278.
- Garland, A. F., Hawley, K. M., Brookman-Frazee, L., & Hurlburt, M. S. (2008). Identifying common elements of evidence-based psychosocial treatments for children's disruptive behavior problems. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*, 505-514.
- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic treatment for antisocial behavior in children and adolescents*. New York: Guilford Press.
- Ingram, B. L., Flannery, D., Elkavich, A., & Rotheram-Borus, M. J. (2008). Common processes in evidence-based adolescent HIV prevention programs. *AIDS and Behavior, 12*(3), 374-383.
- Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36*(4), 567-589.
- Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology, 62*(1), 100-110.
- Kendall, P. C., Flannery-Schroeder, E., Panichelli-Mindel, S. M., Southam-Gerow, M., Henin, A., & Warman, M. (1997). Therapy for youths with anxiety disorders: A second randomized clinical trial. *Journal of Consulting and Clinical Psychology, 65*(3), 366-380.
- Kendall, P. C., Hudson, J., Gosch, E., Flannery-Schroeder, E., & Suveg, C. (2008) Child and family therapy for anxiety-disordered youth: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology 76*, 282-297.

- Kendall, P.C., Kane, M., Howard, B., & Siqueland, L. (1990). *Cognitive-behavioral therapy for anxious children: Treatment manual*. Philadelphia: Department of Psychology, Temple University.
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., . . . Westerberg, H. (2005). Computerized training of working memory in children with ADHD-- A randomized, controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(2), 177-186.
- Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that “everyone has won and all must have prizes”? *Archives of General Psychiatry, 32*, 995–1008.
- National Advisory Mental Health Council Workgroup on Services Research and Clinical Epidemiology (2006). *The road ahead: Research partnerships to transform services*. Washington, DC: Author.
- PracticeWise. (2011). *Evidence-Based Youth Mental Health Services Literature Database*. Satellite Beach, FL: Author.
- Rotheram-Borus, M. J., Swendeman, D., Flannery, D., Rice, E., Adamson, D. M., & Ingram, B. (2009). Common factors in effective HIV prevention programs. *AIDS and Behavior, 13*(3), 399-408.
- Schiffman J., Becker K. D., & Daleiden, E. L. (2006). Evidence-based services in a statewide public mental health system: do the services fit the problems? *Journal of Clinical Child and Adolescent Psychology, 35*, 13-19
- Silverman, W. K., & Hinshaw, S. P. (2008). The second special issue on evidence-based psychosocial treatments for children and adolescents: A 10-year update. *Journal of Clinical*

Child and Adolescent Psychology. Special Issue: Evidence-Based Psychosocial Treatments for Children and Adolescents: A Ten Year Update, 37, 1-7.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore: John Hopkins University Press.

Substance Abuse and Mental Health Services Administration (2006). *From exclusion to belonging: Transforming mental health care in America*. Rockville, MD: Author.

Substance Abuse and Mental Health Services Administration (2011). *National registry of effective programs and practices*. Available on the internet at <http://nrepp.samhsa.gov/>

Task Force on Promotion and Dissemination of Psychological Procedures, Division of Clinical Psychology, American Psychological Association. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist, 48*, 3–23.

Walkup, J. T., Albano, A. M., Piacentini, J., Birmaher, B., Compton, S. N., Sherrill, J. T., . . .

Kendall, P. C. (2008). Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *The New England Journal of Medicine, 359*(26), 2753-2766.

Webster-Stratton, C., & Reid, M. J. (2003). The Incredible Years parents, teachers, and children training series: A multifaceted treatment approach for young children with conduct problems. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 223-241). New York Guilford Press.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55*(4), 542–549.

CHAPTER 3:

Investigating the Fit Between Youths Served by Wraparound Process and Evidence-Based

Treatments

ABSTRACT

This study assessed the degree to which evidence-based treatments (EBTs) fit the problems, demographics, and treatment settings of youths served using wraparound process. Wraparound is a widely implemented and highly popular model for organizing individualized treatments and supports for children with complex needs. A major feature of wraparound is its function of keeping youths in their community and out of restrictive placements via the provision of additional supports. However, wraparound's effectiveness regarding clinical symptom and functioning improvements has long been uncertain. Several investigators have suggested that complementary strengths of wraparound (e.g., real-world transportability, acceptability with diverse stakeholders) and EBTs (e.g., clinical efficacy) may make for a potent combination, especially since wraparound is flexible regarding the specific treatments delivered through its planning process. This study empirically investigated whether EBTs are well-suited for the challenging youths who receive wraparound process. In a large diverse clinical sample, similarities and differences between youths receiving wraparound and non-wraparound services were examined regarding (a) demographic and clinical profiles, (b) "coverability" by treatments in the evidence base, and (c) the nature of practices from the evidence-based that most efficiently serve each group. Results show that coverage for youths receiving wraparound was nearly as high as for youths receiving non-wraparound services. Moreover, the EBT practices identified as most parsimonious for the groups were highly overlapping. These results provide the first large-scale empirical characterization of fit between wraparound and EBTs, and support the proposition that youths receiving wraparound process are well-suited to benefit from EBTs.

INTRODUCTION

Wraparound process (VanDenBerg & Grealish, 1996) has grown over the past two decades to be a major force in child and adolescent mental health. With an estimated 98,000 youths enrolled in over 800 wraparound initiatives across the United States as of 2008 (Bruns, Sather, & Stambaugh, 2008), wraparound may be the nations' most widely implemented community-based model aimed at youth with serious emotional and behavioral health problems (Bruns et al., 2010). This tremendous implementation success is a measure of powerful strengths of the wraparound process with regard to its appeal to diverse stakeholders including families, practitioners, and administrators, as well as judicial and legislative bodies. However, dogged uncertainties regarding the clinical effectiveness of wraparound have long accompanied its success. Wraparound does not dictate the specific clinical practices that are to be implemented via its planning process – on the contrary, individualization is a core tenet of the model – and supportive evidence from controlled trials has been limited (Suter & Bruns, 2009).

Several prominent investigators have noted that this confluence of circumstances makes wraparound a highly promising vehicle for the delivery of evidence-based treatments (EBTs) within systems of care (Weisz, Sandler, Durlak, & Anton, 2006; Tolan & Dodge 2005; Suter & Bruns, 2009; Bruns et al., 2010). The strengths and weaknesses of EBTS – e.g., their considerable documented efficacy but poor implementation track record – seem to be well-suited to compliment wraparound. It is thus envisioned that together wraparound process and EBTs could be a potent combination for reaching children and families with complex needs and alleviating the burden of their mental health concerns. However, some foundational assumptions of this intriguing proposition have yet to be assessed. Little is known about the specific characteristics of youths receiving services via wraparound process. In what ways are they

similar to and different from the participants in research trials, with whom the efficacy of EBTs has been demonstrated? Do EBTs exist in the evidence base that are a good match for youths in need of wraparound process? And if so, do those treatments overlap with other common EBTs, or would making such treatments available be a separate, additional burden for a service system? At the core of these important open questions is whether those served by wraparound process are a “different class of youths” in a sense that meaningfully impacts which clinical treatments are indicated. This paper addresses these issues using relevance mapping (Chopita et al., in press - a), an empirical methodology that compares youths in a given clinical population to the participants from all clinical trials in an evidence base to inform who may be "coverable" by EBTs and which treatments may be most applicable.

Wraparound Process

As stated above, wraparound is an individualized planning process for children and families with complex needs. These needs typically span multiple life areas and the boundaries of traditional, categorical services (VanDenBerg & Grealish, 1996). Often the children in these families are at risk of being removed from the home to residential or institutional treatment. Through additional supports and a family-driven, adaptive planning process that aims to integrate with the child's ecology, wraparound functions to keep youths out of unnecessarily restrictive placements.

Many factors have contributed to wraparound's success. The compelling notion that creating a family- and child-centered approach to services will increase public acceptance and thus implementation of mental health services (Tolan & Dodge, 2005) is one important and perhaps obvious reason. Indeed, wraparound serves a vitally humane function: keeping families together. It is no surprise that in most cases families prefer an option that prevents their child from being

institutionalized or sent far away. Increasingly, legal and governing bodies have mandated the availability of such options, and another force driving the growth of wraparound has thus been large-scale system reform efforts prompted by legislation or class-action lawsuits (Bruns et al., 2010). In the state of California from which the current sample is drawn, both a 1997 senate bill (California Senate Bill 163) and a 2001 legal settlement (*Katie A versus Bonta*) have led to expansion of wraparound, which is now available in most of the state's counties (California Department of Social Services, 2008).

Despite its broad implementation, prominent questions remain about the clinical effectiveness of wraparound process. Allowing youths to stay in their home communities is on its own a very worth outcome. However, wraparound is also intended to improve clinical outcomes, and so far the evidence is slim for wraparound's effectiveness regarding improvements in clinical symptoms and functioning. The wraparound research base has been characterized as "promising" (National Advisory Mental Health Council 2001; New Freedom Commission on Mental Health 2003) or "on the weak side of 'promising'" (Farmer, Dorsey, & Mustillo, 2004, p. 869), and reviewers have often noted weak study designs and an absence of evidence for better outcomes than usual care (Rosenblatt, 1996; Farmer et al, 2004; Weisz et al., 2006). A recent meta-analysis (Suter & Bruns, 2009) identified three experimental studies of wraparound. The study concludes that while these studies yielded a small positive effect when considered together, "better evidence for both the efficacy and the effectiveness of wraparound are sorely needed" (p. 347). Indeed, taken individually, the experimental studies reviewed produced largely non-significant findings regarding their respective targeted outcome symptoms (Carney & Buttell, 2003; Clark, Lee, Prange, & McDonald, 1996; Evans, Armstrong, & Kuppinger, 1996). Also notable are the unique challenges to the evaluation of wraparound

including the difficulty of separating the effects of specific services provided from the encompassing wraparound process, and the variability between wraparound programs given the absence of a universally adopted manual (Suter & Bruns, 2009).

Regardless of the effectiveness labels, the feasibility, acceptability, and generalizability of wraparound demonstrated by its wide implementation mean that this model cannot be ignored by researchers or practice systems. Moreover such stakeholders have a lot to gain from embracing wraparound's tremendous strengths. The real-world transportability of wraparound is in stark contrast to that of standard evidence-based treatments to date (Hoagwood et al., 2001; Weisz et al., 2006) whose huge federal research investment and low implementation payoff may be compared to a Ferrari idling in the nation's garage with no wheels.

Several investigators have suggested that wraparound may be able to provide “wheels” for evidence-based practices, and that EBTs may in turn be able to offer an answer to the clinical effectiveness challenges that wraparound has faced (Weisz et al., 2006; Tolan & Dodge 2005; Suter & Bruns, 2009; Bruns et al., 2010). One reason this seems feasible is that wraparound is flexible regarding the specific clinical practices delivered. Though the practice model for wraparound has evolved considerably (Walker & Bruns, 2006), it is important to emphasize that wraparound is not a clinical treatment. Rather, Bruns & colleagues (2010) explain that the availability of effective treatments is a necessary condition for wraparound to be successful: "the behavioral health system must be able to provide wraparound-enrolled youth and families with access to effective treatments and ancillary supports. Without access to a range of effective clinical treatments and supports ... wraparound teams will find it more difficult to effectively strategize to meet the full range of youth and family needs" (p. 316). Indeed, wraparound's “plan development” phase focuses on discussion of treatments and strategies that have been successful

in the past (Bruns et al., 2010), and information from the general services literature on EBTs may make a valuable contribution by helping families to understand which choices are most likely to help their children. But despite these suggestions in the literature, so far the idea of combining wraparound and EBTs has received little direct investigation.

The Present Study

This paper's purpose is to empirically investigate the feasibility of delivering EBTs via wraparound. Specifically, this study focused on whether the treatments available in the evidence base are a good match for the youths served by wraparound. We approach this question by examining the demographic and clinical profiles of youths served with wraparound, and comparing them to youths receiving standard services and also to participants in the randomized clinical trials (RCTs) that provide the evidence for EBTs. Moreover, the analysis asks what part of the clinical content in the evidence base may be most appropriate for wraparound, and how do the practices indicated for youths served by wraparound compare to those indicated for youths receiving standard services.

To investigate these questions, we use relevance mapping, an analytic framework for coordination of evidence-based treatments (Chorpita, Bernstein, & Daleiden, in press). Relevance mapping involves structured comparison of client characteristics in a service sample to participant characteristics from studies of EBTs. By checking each child in the sample against each EBT study, relevance mapping results can reveal who is "coverable" by any EBT, under different definitions of matches between study participants and clients. Further, relevance mapping can identify minimum sets of treatments needed to serve a group of clients. In this study we compare both of these types of results for youths receiving wraparound and standard services.

In summary, wraparound is an appealing overall approach to service delivery. It is much more widely used than EBTs (Suter & Bruns, 2009), has a much stronger record of engaging family and community support, and unlike EBTs it addresses the organization of multiple interventions and services (Weisz et al., 2006). And importantly, the specific services offered within the context of wraparound are free to vary with available services in the community (Weisz et al., 2006; Suter & Bruns, 2009). As for EBTs, there are many indications that their proponents need to get better at finding creative ways to reach consumers (e.g., Swendeman & Rotheram-Borus, 2010; Kazdin & Blase, 2011). These factors combine to suggest that wraparound and EBTs may make a potent team. However, no data is yet available regarding whether EBTs are actually well-suited for the challenging youths who typically receive wraparound process. This study therefore investigated this question, specifically examining the similarities and differences between youths receiving wraparound and non-wraparound services with regard to (a) demographic and clinical profiles, (b) coverability by the treatments in the evidence base, and (c) the nature of the practices from the evidence-based that most efficiently serve each group. Additionally, discussed throughout are the situations where changes to treatment setting or adaptations and extension of existing treatments may be needed for youths in this study's wraparound and standard services groups to get the most benefit from evidence-based practices.

METHOD

Client Sample

The client data sample for this study comes from EMQ FamiliesFirst (EMQFF), a large mental health service agency with a service population well distributed between northern, southern and central California. This sample represents 3,932 youths receiving services from

EMQFF between January 2009 and May 2010. The only inclusion criteria were data availability and age less than 19 years. To be included, a child's episode of care had to have data available on primary problem, age, gender, and setting. The data availability criterion was used because these variables were required in the analyses considered in this report, and so otherwise the applicability of the research findings to a particular child could not be estimated. For youth with multiple episodes of care that met these criteria, only the first episode during this time period was considered.

The resulting full client sample was then divided into two subsamples: youths receiving wraparound services (Wrap) and youths receiving services not through the wraparound model (Non-Wrap). Youths were allocated to the appropriate subsample based on the episode of care selected as described above. Table 1 reports demographic characteristics of the youths in the full sample along with the Wrap ($n = 828$) and Non-Wrap ($n = 3,104$) subsamples. The mean age for the Wrap group was 13.90 ($SD = 3.49$), and for the Non-Wrap group was 11.76 ($SD = 4.14$).

Chart diagnoses were based on Diagnostic and Statistical Manual of Mental Disorders (4th ed. [DSM-IV]; American Psychiatric Association, 2000) codes. For the purposes of this investigation, DSM-IV diagnoses present in the charts were cross-walked to 10 broad categories (the mapping of all diagnoses to problem groups is available upon request), shown in Table 1 along with their frequencies in this population. The mean number of diagnoses for the Wrap group was 1.81 ($SD = 1.15$), and for the Non-Wrap group was 1.57 ($SD = 0.86$). The decision to use a small number of broad problem categories was made to provide a common framework for matching clients to studies, given that research studies use a wide variety of taxonomies and methods for classifying and enrolling participants, not limited to diagnosis. Chorpita et al. (in press - a) describe the tradeoff involved in this decision between the precision with which a

problem is characterized and the availability of data to inform a match. Given that less than half the studies in our review reported clinical diagnosis, we deemed the data availability concern to be more pressing and thus opted for the broad categories.

Another major parameter considered from the client dataset is the setting in which youths received treatment (e.g., clinic, day care, residential). EMQFF provides services via numerous distinct treatment programs each associated with a particular setting. The settings of these programs were mapped to three broad categories to provide a common framework for matching clients to studies: “clinic,” “home, school, or community-based,” and “residential.” These broad categories were selected to represent what may be common assumptions about the setting considerations most important for treatment generalizability. The decision to use a small number of broad setting categories was made after considering an alternate setting definition including 18 categories (e.g., “corrections,” “hospital,” “summer day camp”). As with the primary problem parameter, in characterizing setting a tradeoff is involved between the precision with which treatment settings are described and the availability of data to inform a match. Although inferences about treatments provided in more specific setting categories (e.g., “foster home”) can perhaps be made with greater confidence than is possible with the broader categories (e.g., “residential”), the literature relevant to those settings, and hence the EBTs available to match to clients in a given category, is greatly reduced in this tradeoff.

Study Dataset

The study dataset involved codes from 524 RCTs of child mental health treatments. In order to determine which treatments are candidates for analysis, relevance mapping requires the user to define or select a standard of evidence that identifies which treatments in the study dataset are considered evidence-based. We employed a standard of evidence-based largely on the criteria

developed and employed by the Task Force on Promotion and Dissemination of Psychological Procedures (1995). According to this standard, a manualized treatment must (a) show statistically superior outcomes to a waitlist or no-treatment control group in at least two randomized trials, (b) show statistically superior outcomes to an active treatment or psychological placebo in at least one randomized trial, or (c) show equivalent outcomes to an already established evidence-based treatment in at least one randomized trial in which the average group size is at least 30 participants (see Chorpita et al., in press - b, for additional details). This standard was met by 288 protocols tested with 318 participant groups in 255 RCTs. The problem areas targeted by each RCT and the settings in which treatments were provided were reduced to the same broad categories used with the client dataset. The studies with treatments meeting the evidence criteria corresponded to the following problem areas: anxiety (n studies = 75), attention/hyperactivity (n studies = 24), autism spectrum (n studies = 10), depression (n studies = 23), disruptive behavior (n studies = 89), eating disorders (n studies = 4), substance use (n studies = 19), and traumatic stress (n studies = 9) with two studies not mapping to any of these categories. Among the 318 participant groups, 239 had treatment setting information available (setting information was not reported in published research papers for the remainder of groups; some groups that did have information available received treatment in multiple settings), and these corresponded to the following settings: clinic (n groups = 110), home, school, or community based (n groups = 116), and residential (n groups = 23), with one group not mapping to any of the three categories. Additional characteristics of a recent version of this dataset have been reported in previous research (Chorpita et al., in press - b).

Treatment practices, which we refer to as practice elements, were coded from each treatment in the study dataset. Practice elements are discrete procedures that are structured components of

a larger course of treatment (Chorpita, et al., 2005; Chorpita & Daleiden, 2009). Examples of practice elements include "Time Out," "Relaxation Training," and "Psychoeducation." Two raters coded each protocol regarding the presence or absence of 59 PEs, and an expert rater performed a final validation and review of all codes. A detailed description of the PE coding and reliability is reported by Chorpita & Daleiden (2009).

Procedure

Relevance mapping. The analytic procedures for this study followed the relevance mapping framework, described in detail in a previous paper (Chorpita et al., in press - a). The central part of the relevance mapping model involves taking every client in a given service population and determining which published research trials have participants with matching characteristics. Each child in the client data sample is compared to each study in a study dataset to determine if the child is “covered” by that study. These structured comparisons can use any variables common to the client and study datasets, and in the current study, primary problem, age, gender, and setting were used for matching children to studies (see *Data Analysis*). When a client is found to match a particular study on these parameters, a record of the client-study combination is added to a list of all matches. The resulting list of matches between children and studies is then summarized to answer questions about the characteristics of coverable and non-coverable youth.

Additionally, the list of matches is the starting point for a further set of optimization analyses that aim to find the smallest sets of treatments or treatment components that combine to cover the most youth. These best complimentary combinations are found by an automated search that tests many different groupings of treatments to find the smallest sets that have matches in the list for a criterion percentage of youth. Details of the matching and the optimal treatment set identification analyses have been reported previously (Chorpita et al., in press - a).

Data Analysis

A first set of analyses were used to investigate whether there were significant differences between the Wrap and Non-Wrap groups on the variables considered for matching youths to treatments in this study. Chi-square tests examined the relationships between wraparound status and primary problem (9 tests) as well as gender (1 test). Additionally, group differences were examined in mean age and number of diagnoses. Since preliminary tests (i.e., Levene's test of homogeneity of variance) indicated heterogeneous variance between the groups for both comparisons ($p < .001$), Welch's t test (which is robust to heterogeneity of variance) was employed. All analyses were performed using a 99.6% confidence interval (alpha of .004 after a Bonferroni correction for the 12 tests performed using a 95% confidence interval). Tests were omitted for the substance use and eating problem areas since cells contained less than five participants, violating the assumptions of the chi-square independence test. Similarly, differences in setting were not assayed using tests of statistical significance since the mapping of all wraparound youths to the category "home, school, or community based" made setting differences obvious.

All other analyses utilized the relevance mapping model. Two scenarios were analyzed, representing combinations of parameters from the study dataset and the client dataset on which clients and research participants must match: problem-age-gender (PAG) and problem-age-gender-setting (PAGS). The PAG scenario defines an EBT as relevant to clients with the same primary problem, within the same age range, and with the same gender as the participants in a study in which that treatment was successfully tested,⁶ and the PAGS scenario additionally

⁶ For the "problem" parameter, coverage required the problem area targeted for treatment in a given study and used for participant inclusion in that study to match the primary problem area of a child in the client dataset. For example, a client with a primary problem area of depression could only be covered by treatments from studies that targeted

requires a match between the setting where study participants received and the setting where the client received treatment. These scenarios were selected to represent what may be common assumptions about the parameters most important for treatment generalizability. The PAGES scenario was determined *a priori* to be the primary scenario of interest given the importance of setting considerations to understanding the goodness of fit of EBTs to youths receiving wraparound. For this reason, only the PAGES scenario was used for identifying best sets of treatment practices. However, given the particular importance of setting as a consideration regarding the fit of EBTs for youths receiving wraparound process, the PAG scenario was included in the analysis of coverability so as to provide a reference point, making visible the specific effects of setting as a matching requirement for the two groups.

For both groups, the total number of youths coverable was calculated for the PAG and PAGES scenarios, and characteristics of non-coverable youths were examined by counting the number of non-coverable youths with each level of the problem, age, gender, and setting variables. Best sets of practice elements were then analyzed under the PAGES scenario, with separate searches performed for the Wrap and Non-Wrap groups. For each practice element in the best sets, two additional statistics were calculated: (a) the percentage of youths to whom the practice element applied, and (b) the percentage of youths who would no longer be coverable were the practice element to be dropped from the identified service array.

RESULTS

Table 1 shows the results of the analyses comparing the Wrap and Non-Wrap groups on the diagnostic and demographic variables, with a superscript in the row labels marking categories where significant differences were found. Significantly more youths receiving wraparound

the treatment of depression and included youth with depression as a primary problem. For age and gender, to cover a given client a study was required to have at least one participant with matching values.

services had primary problems of traumatic stress (11% vs. 6%), mania (8% vs. 3%), and no diagnosis (6% vs. < 1%). Significantly less youths receiving wraparound services had primary problems of depression (20% vs. 29%) and primary problems that fell into the “other problems” category (20% vs. 28%). Further exploring the “other problems” category, examination of diagnoses indicated that the seven most common diagnoses for youths in the “other problems” category were the same for the Wrap and Non-Wrap groups.⁷ Significant differences were also found for mean number of diagnoses, $t(1083.14) = 31.60, p < .001$, and mean age, $t(1508.68) = 226.21, p < .001$.

As noted above, no significance tests were used for the setting categories since the mapping of all youths in the Wrap group to the category "home, school, or community based" made setting differences obvious. For the Non-Wrap group, "home, school, or community based" was also the most common setting category, representing 75% of the group's youths. For the Non-Wrap group, 23% received services in the clinic setting, while 2% received residential treatment.

The coverage results are also shown in Table 1. Coverage for the Wrap and Non-Wrap groups was analyzed under both the PAG and PAGES scenarios. Overall, coverage was similar but slightly lower for Wrap than Non-Wrap under both the PAG (41% and 35% not covered, respectively) and PAGES (42% and 38% not covered, respectively) scenarios. For both groups, the two most common problem categories – depression and disruptive behavior – were almost entirely covered under both scenarios. The largest category of uncovered youth for both groups was “other problems.” These youths by definition could not be covered by any treatment since matching on problem was required.

⁷ The most common diagnoses mapped to the “other problems” category (with n's listed in parentheses for Wrap and Non-Wrap, respectively) were: adjustment disorders (n =51, 255), other mood disorders (n =51, 208), anxiety disorder not otherwise specified (NOS; n =19, 141), disorders of infancy or early childhood NOS (n = 6, 120), other psychotic disorders (n = 9, 55), intermittent explosive disorder (n = 9, 42), and reactive attachment disorder (n = 14, 26).

Amidst the patter of largely similar results for the two groups, notable differences in coverage can also be seen in Table 1. First, there is a large decrease in coverage for Non-Wrap youths in the residential setting under PAGES. This decrease can be conceptualized as a *main effect* of residential within the PAGES scenario. It is likely that many of these youths would be coverable if treated in a home, school, or community based setting such as that used with wraparound. Other coverage differences relate to the *interaction* of setting with the various combinations of problem, age, and gender. For example, in the Non-Wrap group coverage decreases when moving from the PAG scenario to PAGES for the problem categories attention (8% decrease) and traumatic stress (16% decrease), but comparable drops in coverage were not found for Wrap. Further investigation revealed that 84% of the decrease in coverage for attention and 33% for traumatic stress involved youths in the clinic and residential settings. Most of these youths would have been found coverable had they been receiving wraparound or other services in a home, school, or community based setting. Conversely, 67% of the decrease for traumatic stress involved youths in the home, school, or community based setting, and most of those youths would have been found coverable had they been receiving treatment in the clinic setting. A similar story is true for youths ages 4 to 6. Table 1 shows that the decrease in coverage when moving from PAG to PAGES for Wrap (15% decrease) was approximately double that found for Non-Wrap (8% decrease). Further investigation revealed that half of the decrease in coverage was due to youths in the clinic setting and half was due to youths in the home, school, or community based setting. Each of these two clusters of “uncoverable” youths would likely have been found coverable if they had received treatment in the other clusters’ setting.

Other interesting coverage differences can be seen that do not relate to setting. For example, autism spectrum coverage was considerably worse for Wrap than for Non-Wrap. However, since

the portion of youths uncovered was the same in PAG and PAGES, setting was not involved in the disparity. Further investigation revealed age to be the primary factor involved, as the uncovered youths were almost entirely above age 13 with few treatments shown effective for that age group. Wrap coverage was worse for this area simply because the autism spectrum youths receiving wraparound process were mostly above age 13.

The results of the best set identification analyses are shown in Tables 2 and 3. These tables list the smallest possible groupings of practice elements that combine to cover 100% (Table 2) and 98% (Table 3) of all possible youths. 38 practice elements were required to cover the full 480 coverable youths in the Wrap group, and 41 were needed for the 1,920 youths coverable in the Non-Wrap group. The number of practice elements needed decreased considerably when the coverage requirement was reduced to 98% of coverable cases: 33 practice elements were sufficient for Wrap, and 30 for Non-Wrap.

The first big-picture result to note from the practice element rows of these tables is that the profile of practices identified for Wrap and Non-Wrap is largely overlapping. For the 100% coverage requirement (Table 2), all but one of the practice elements identified for Wrap was also identified for Non-Wrap. At 98%, 23 of the 33 practice elements identified for Wrap were also identified for Non-Wrap, and all of the remaining 10 were identified for Non-Wrap when 100% coverage was required (Table 2), indicating that they would also be useful for providing non-wraparound services.

The two columns listing percentages for each group provide additional information about the similarities and differences for Wrap and Non-Wrap. The rightmost column for each group refers to the percentage of youths who would no longer be coverable were a particular approach to be dropped from the identified service array. The tables thus show that the top three practice

elements uniquely required for Wrap and Non-Wrap (at both 100% and 98%) were the same: psychoeducation-child, maintenance/relapse prevention, and problem solving. The Non-Wrap group has more practice elements with low values (i.e., < 1%) in this column in the 100% coverage table (Table 2), which corresponds to the greater decrease in number of practice elements required for the Non-Wrap group at 98% (30, down from 41) compared to the Wrap group (33, down from 38). This is an indicator of greater heterogeneity in the Non-Wrap group, since many of the identified practices were only required to serve a small number of the group's youths. The middle column for each group shows the total number of youths to whom each practice element was applicable. These percentages are considerably higher for Wrap than for Non-Wrap in general, again indicating greater heterogeneity for the Non-Wrap group. However, for both Wrap and Non-Wrap, all identified practice elements apply to a much higher percentage of youths than would be lost to coverage if the practice were dropped from the service array, indicating considerable redundancy within the minimal sets. In other words, for both Wrap and Non-Wrap, most youths were covered by multiple groupings of practice elements from among the identified service arrays, and a service system implementing the full identified array would thus have various options when treating those youths.

DISCUSSION

This study assessed the degree to which evidence-based practices fit the problems, demographics, and treatment settings of youths served using wraparound process. The results show that coverage for youths receiving wraparound was nearly as high as for the youths receiving non-wraparound services. Moreover, the evidence-based practices identified as most parsimonious for the Wrap group were almost entirely overlapping with those identified for the Non-Wrap group. These results provide the first large scale empirical characterization of the fit

between wraparound and EBTs. This evidence supports the proposition that youths receiving wraparound process are not an altogether different class of children, with clinical problems to which the mental health evidence base does not apply. Rather, at least within the large and diverse community sample examined for this study, EBTs appear to be a good fit for these youths, and wraparound appears to be a very promising vehicle for the delivery of EBTs.

The results also provide a compelling argument that wraparound process brings some children closer to a context where they may receive greater benefit from evidence-based practices. The results section details several examples where this may be the case, but the most salient example – youths receiving services in residential settings – is also the most in line with core motivations of wraparound process. While this category represented only a small portion of the overall sample, 83% of such youths were not coverable when matching on setting was required. These youths may be getting “left behind” from the most appropriate EBTs by remaining in residential treatment. Not only could wraparound process enable less restrictive care for such children, it may also enable them to move to a context where they could benefit more from EBTs.

A number of significant differences were found in the prevalence of primary problem areas between the groups, and in some cases (e.g., mania) this contributed to the somewhat lower degree of coverage for Wrap. Additionally, Wrap youths were found to have more diagnoses on average than Non-Wrap youths. While that result presents no surprise given wraparound’s function of serving challenging cases, the degree of the difference – just 0.24 diagnoses more on average – may be less than anticipated.

More interesting is that in spite of the differences identified between youths in the two groups, youths receiving wraparound largely matched to the same EBTs as those receiving

standard services. This finding in no way refutes a core purpose of wraparound, which is to provide extra supports to youths for whom standard services alone would likely be insufficient and lead to escalating levels of restrictive care. Rather, these results characterize the nature of the clinical procedures that may be most helpful for these youth when delivered *along with* the extra supports that wraparound provides and the family-driven planning in which it excels. In other words, receiving the additional benefits of wraparound should not preclude youths from also having available the best treatments that match their clinical presentation.

Similarly, it is important to emphasize that this paper in no way advocates against the core wraparound values of family and youth choice in the treatment planning process. The results of a local relevance mapping analysis should not be used by professionals to singlehandedly prescribe a treatment plan. Rather, such information should be used to inform the individualized planning process as well as the training of wraparound professionals, such that they can offer families a choice among options that have strong evidence of working with children most like their own. The availability of multiple choices afforded by the practice elements in the identified service arrays fits well with these wraparound values. Moreover, redundancy in coverage is often a practical necessity given issues faced by all service systems such as considerable failure rates among even the best treatments and the frequent need for additional services even after a “successful” first treatment episode (cf. Hansen, Lambert, & Forman, 2002).

For EBTs to be made most effective for youths receiving wraparound services, it will be important to consider whether any changes to the EBTs themselves are needed. This paper’s results indicate that in general EBTs are a good match for the youths served with wraparound in the current sample, even when only considering treatments shown to work in a setting like the community context where much of wraparound takes place. For this majority of coverable

youths, relatively few changes may thus be needed to make EBTs work. However, for both the Wrap and Non-Wrap groups, many youths were not coverable by anything in the evidence base (compare similar findings from Chorpita et al., in press - a; Bernstein, Chorpita, Daleiden, Ebesutani, & Rosenblatt, 2011). Chorpita et al. (in press - a) describe three ways to proceed in such cases: redirection, adaptation, and extension. The results section above documents evidence that in several circumstances common to the current sample, changing the treatment setting (i.e., *redirection*) both to and from wraparound's community context could put children in a setting where they may receive greater benefit from evidence-based practices. The other two options, adaptation and extension, are also important since redirection is not always possible or desirable. Practices that are found to be a good fit under less restrictive requirements (like PAG or even a scenario requiring matching on primary problem only) can be *adapted* to be age, gender, or setting appropriate or used unmodified (i.e., *extended*) for youths receiving wraparound. Conversely, as was noted in the results, adaptations or extensions are also warranted in many circumstances to apply EBTs to the clinic or residential settings for youths receiving standard services. Regarding adaptation, one might ask what will be needed to make cognitive behavioral therapy, for example, most effective in wraparound. While answers to many such questions remain to be investigated, encouraging evidence already exists from researchers like Kolko et al. (2009) who found a disruptive behavior treatment originally developed for a clinic setting to produce comparable improvements when modularized and applied in a community context.

The limitations of the current findings are also important to clarify. Two notable issues relate to the definitions of the setting and problem factors used for matching children to treatments. Regarding settings, we mapped wraparound to the setting category "home, school, or community

based,” reflecting an assumption that the home and school contexts are similar enough to community-based treatments that mutual coverage should be afforded. Bruns and colleagues (2010) note that, “After *family and youth voice and choice*, perhaps the most important and enduring principles of wraparound are those of *individualized* and *community-based*” (p. 315). Services delivered through wraparound can be provided in the home or school (indeed in many cases they are delivered in multiple such settings), and arguably a broad definition of community-based is inclusive of these settings. As noted above (see *Methods*) this decision was made after considering the tradeoff involved between the precision with which treatment settings are described and the availability of data to inform a match. Choosing a more granular approach to defining setting would also be defensible, and would decrease coverage for both the Wrap and Non-Wrap groups. Future research should directly examine the effects of varying definitions of setting on coverage for youths receiving wraparound and standard services.

Regarding definition of the problem factor, we configured the analysis to match on youths’ primary problems only (i.e., those derived from their primary DSM-IV diagnoses), ignoring comorbid problems for the purposes of matching. We made this choice because EBTs have generally been designed and tested for a single primary problem. However, as with setting, stricter definitions of problem would also be defensible and would considerably decrease coverage for both groups. In the case of the problem factor, a definition requiring matching on comorbid problems may be expected to cause a somewhat greater decrease in coverage for the Wrap group, since the mean number of comorbid diagnoses was slightly higher.

It is also important to clarify that much is needed beyond EBTs to make wraparound successful. Indeed, most of the well-known factors from the burgeoning literature regarding EBT implementation apply to wraparound (e.g., training, coaching, administrative supports, feedback;

Fixsen et al., 2005; Glisson et al., 2008), along with additional wraparound-specific requirements such as the flexible fiscal structures needed to rapidly purchase diverse services for the individual needs of wraparound youths (Bruns et al., 2010). Bruns and colleagues (2010) explain that “current conceptualizations of wraparound include an implementation ‘blueprint’ that specifies a set of key areas in which system- and program-level structures and procedures must be established” (p. 316). Thus EBT service arrays like those identified in the current analysis are just one component of what is required to make wraparound most effective, and they speak specifically to the clinical procedures that best fit the needs of the youths wraparound serves.

Finally, while the identity of the specific practice elements listed in this paper’s results is not of general significance (since relevance mapping is an inherently local methodology; Chorpita et al., in press - a), the concept demonstrated may have important implications for practice organizations regarding how to select efficient treatment arrays to serve diverse populations. The current paper adds to the evidence that knowing the relevance of a treatment to a specific clinical population is an important aspect of building a comprehensive service array in a local context. In practice, the 30 to 41 common elements identified in our analysis may be more than a service organization could feasibly implement. For an organization pursuing implementation, it is valuable to know what is required to serve 100% and 98% of coverable youths, but the final determination of which practices to make available would also involve consideration of fiscal, administrative, and workforce capacity restrictions. For example, an organization may determine that its capacity is 20 practice elements; a further relevance mapping search could then be performed to find the best set of 20 that together cover as many youths as possible from the organization's service sample. Bernstein et al. (2011) describe additional options whereby practice elements can be combined with traditionally packaged EBT programs (e.g.,

Multisystemic Therapy; Henggeler et al., 1998) to take advantage of the treatments already available in a clinical workforce. This paper illustrates the first steps of this type of decision process for an organization needing to balance large portions of youths receiving both wraparound and standard services.

In summary, this paper's central findings provide evidence that EBTs are a good fit for youths served by the wraparound process. Wraparound's tremendous availability, acceptability (to families, communities, and practitioners alike), and values-based approach to treatment planning thus make it a highly promising vehicle to increase the reach of treatments with strong demonstrated efficacy that have to date reached few consumers. EBTs likewise have great potential to increase the effectiveness of clinical outcomes for wraparound, providing an answer to a central wraparound critique, while ameliorating mental illness burden for the youths served by this popular process.

.

Table 1.

Youths diagnostic and demographic characteristics and percentage of children in the corresponding categories not coverable (NC) by evidence-based treatments identified in published randomized clinical trials (N studies = 524).

| | N | | | Scenario | | | |
|-------------------------------|-------|------|----------|-----------|----------|------------|----------|
| | Total | Wrap | Non-Wrap | PAG (%NC) | | PAGS (%NC) | |
| | | | | Wrap | Non-Wrap | Wrap | Non-Wrap |
| Total | 3,932 | 828 | 3,104 | 40.6% | 35.3% | 42.0% | 38.1% |
| Problem type | | | | | | | |
| Depression ^a | 1,068 | 162 | 906 | 1% | 2% | 1% | 4% |
| Disruptive behavior | | | | | | | |
| Attention | 396 | 95 | 301 | 49% | 30% | 51% | 38% |
| Traumatic stress ^a | 276 | 95 | 181 | 0% | 1% | 6% | 17% |
| Mania ^a | 150 | 67 | 83 | 94% | 86% | 100% | 95% |
| Anxiety | 105 | 17 | 88 | 0% | 2% | 0% | 5% |
| Autism spectrum | 68 | 9 | 59 | 78% | 34% | 78% | 42% |
| Substance use | 5 | 4 | 1 | 0% | 0% | 0% | 0% |
| Eating | 2 | 1 | 1 | 0% | 0% | 100% | 0% |
| None ^a | 63 | 52 | 11 | 100% | 100% | 100% | 100% |
| Other problems ^a | 1,043 | 165 | 878 | 100% | 100% | 100% | 100% |
| Age ^a | | | | | | | |
| 0 to 3 | 108 | 4 | 104 | 100% | 75% | 100% | 83% |
| 4 to 6 | 382 | 41 | 341 | 56% | 55% | 71% | 63% |
| 7 to 9 | 521 | 71 | 450 | 39% | 34% | 42% | 37% |
| 10 to 12 | 702 | 78 | 624 | 28% | 27% | 32% | 31% |
| 13 to 15 | 1,237 | 296 | 941 | 36% | 31% | 37% | 32% |
| 16 to 18 | 982 | 338 | 644 | 45% | 33% | 45% | 34% |
| Gender ^a | | | | | | | |

| | | | | | | | |
|-------------------------------------|-------|-----|-------|-----|-----|-----|-----|
| Boys | 2,280 | 523 | 1,757 | 42% | 36% | 44% | 39% |
| Girls | 1,652 | 305 | 1,347 | 38% | 34% | 39% | 37% |
| Setting ^b | | | | | | | |
| Home, school, or community based | 3,155 | 828 | 2,327 | 41% | 36% | 42% | 37% |
| Clinic | 713 | 0 | 713 | NA | 33% | NA | 36% |
| Residential | 64 | 0 | 64 | NA | 33% | NA | 83% |

Note. Scenario column headings refer to parameters on which clients and research participants must match. PAG = primary problem-age-gender; PAGES = primary problem-age-gender-setting. Ns refer to the total youths within the row label class; percentages refer to the percent of those youths not coverable.

^a Significant difference in the proportion of Wrap and Other youth. Significance was set at $p = .004$ after a Bonferroni correction for the 12 tests performed using a 95% confidence interval.

^b No significance tests were applied to setting categories since all youths in the Wrap group mapped to the category “home, school, or community based.”

Table 2.

Practice Elements (PEs) relevant to youths in the PAGES scenario with coverage criterion of 100% of coverable youth.

| Cases Coverable Minimum PEs | Wrap | | | Non-Wrap | | |
|----------------------------------|--------------------------|--|---|------------------------------|--|---|
| | 480 (58% of Wrap cases) | | | 1920 (62% of Non-Wrap cases) | | |
| | 38 | | | 41 | | |
| PE | PE in minimum set? | Applies to this % of sample youth | % of youths lost if PE dropped | PE in minimum set? | Applies to this % of sample youth | % of youths lost if PE dropped |
| Psychoeducation-Child | ✓ | 44.8% | 20.8% | ✓ | 47.9% | 17.8% |
| Maintenance/Relapse Prevention | ✓ | 49.8% | 13.6% | ✓ | 52.7% | 14.0% |
| Problem Solving | ✓ | 54.0% | 11.5% | ✓ | 56.0% | 6.0% |
| Insight Building | ✓ | 52.5% | 8.9% | ✓ | 44.9% | 2.5% |
| Communication Skills | ✓ | 38.9% | 7.1% | ✓ | 47.7% | 3.7% |
| Cognitive | ✓ | 48.6% | 5.3% | ✓ | 51.4% | 3.9% |
| Relaxation | ✓ | 39.7% | 4.3% | ✓ | 46.8% | 1.8% |
| Emotional Processing | ✓ | 25.4% | 3.7% | ✓ | 21.0% | 0.4% |
| Psychoeducation-Parent | ✓ | 46.3% | 3.0% | ✓ | 56.5% | 4.2% |
| Tangible Rewards | ✓ | 43.7% | 3.0% | ✓ | 49.1% | 3.0% |
| Praise | ✓ | 25.4% | 2.9% | ✓ | 28.2% | 3.5% |
| Relationship/Rapport Building | ✓ | 38.3% | 2.8% | ✓ | 39.6% | 2.0% |
| Response Cost | ✓ | 24.9% | 2.8% | ✓ | 23.3% | 1.9% |
| Monitoring | ✓ | 22.8% | 2.5% | ✓ | 24.2% | 1.5% |
| Modeling | ✓ | 16.1% | 1.7% | ✓ | 32.3% | 2.6% |
| Therapist Praise/Rewards | ✓ | 21.0% | 1.6% | ✓ | 28.5% | 1.7% |
| Talent or Skill Building | ✓ | 21.4% | 1.6% | ✓ | 21.2% | 1.6% |
| Goal Setting | ✓ | 40.0% | 1.6% | ✓ | 48.0% | 1.0% |
| Social Skills Training | ✓ | 45.5% | 1.3% | ✓ | 52.7% | 1.3% |
| Parent Coping | ✓ | 38.2% | 1.3% | ✓ | 40.3% | 1.0% |
| Natural and Logical Consequences | ✓ | 19.4% | 1.3% | ✓ | 21.0% | 0.7% |

| | | | | | | |
|---|---|-------|------|---|-------|------|
| Educational Support | ✓ | 21.5% | 1.3% | ✓ | 19.2% | 0.6% |
| Crisis Management | ✓ | 37.0% | 1.3% | ✓ | 33.2% | 0.5% |
| Family Engagement | ✓ | 19.3% | 1.3% | ✓ | 11.0% | 0.5% |
| Family Therapy | ✓ | 18.7% | 1.3% | ✓ | 16.0% | 0.5% |
| Functional Analysis | ✓ | 18.7% | 1.3% | ✓ | 10.2% | 0.5% |
| Individual Therapy for Caretaker | ✓ | 18.2% | 1.3% | ✓ | 11.0% | 0.5% |
| Marital Therapy | ✓ | 18.2% | 1.3% | ✓ | 10.2% | 0.5% |
| Self-Monitoring | ✓ | 43.7% | 1.2% | ✓ | 47.6% | 3.4% |
| Peer Pairing | ✓ | 22.1% | 1.2% | ✓ | 17.8% | 0.8% |
| Exposure | ✓ | 10.1% | 1.0% | ✓ | 10.2% | 3.5% |
| Motivational Interviewing | ✓ | 15.0% | 0.5% | | | |
| Self-Reward/Self-Praise | ✓ | 25.2% | 0.4% | ✓ | 36.5% | 0.4% |
| Activity Scheduling | ✓ | 19.8% | 0.2% | ✓ | 29.4% | 1.5% |
| Attending | ✓ | 3.0% | 0.1% | ✓ | 12.4% | 1.2% |
| Time Out | ✓ | 3.0% | 0.1% | ✓ | 13.3% | 1.0% |
| Commands | ✓ | 3.7% | 0.1% | ✓ | 12.5% | 0.9% |
| Differential Reinforcement | ✓ | 3.7% | 0.1% | ✓ | 14.1% | 0.9% |
| Behavioral Contracting | | | | ✓ | 33.9% | 0.4% |
| Assertiveness Training | | | | ✓ | 16.4% | 0.3% |
| Biofeedback/Neurofeedback | | | | ✓ | 3.7% | 0.2% |
| Stimulus Control or Antecedent Management | | | | ✓ | 18.1% | 0.2% |

Table 3.

Practice Elements (PEs) relevant to youths in the PAGES scenario with coverage criterion of 98% of coverable youth.

| Cases Coverable Minimum PEs | Wrap | | | Non-Wrap | | |
|--------------------------------|--------------------------|--|---|------------------------------|--|---|
| | 480 (58% of Wrap cases) | | | 1920 (62% of Non-Wrap cases) | | |
| | 33 | | | 30 | | |
| PE | PE in minimum set? | Applies to this % of sample youth | % of youths lost if PE dropped | PE in minimum set? | Applies to this % of sample youth | % of youths lost if PE dropped |
| Psychoeducation-Child | ✓ | 44.8% | 19.4% | ✓ | 47.9% | 18.9% |
| Maintenance/Relapse Prevention | ✓ | 49.8% | 13.3% | ✓ | 52.7% | 14.2% |
| Problem Solving | ✓ | 54.0% | 11.5% | ✓ | 56.0% | 12.4% |
| Insight Building | ✓ | 52.5% | 11.5% | ✓ | 44.9% | 6.0% |
| Communication Skills | ✓ | 38.9% | 7.6% | ✓ | 47.7% | 3.2% |
| Cognitive | ✓ | 48.6% | 5.2% | ✓ | 51.4% | 11.0% |
| Relaxation | ✓ | 39.7% | 4.3% | ✓ | 46.8% | 6.4% |
| Social Skills Training | ✓ | 45.5% | 4.1% | ✓ | 52.7% | 1.4% |
| Emotional Processing | ✓ | 25.4% | 3.7% | B | 21.0% | 0.4% |
| Relationship/Rapport Building | ✓ | 38.3% | 3.5% | ✓ | 39.6% | 5.7% |
| Tangible Rewards | ✓ | 43.7% | 3.0% | ✓ | 49.1% | 5.4% |
| Psychoeducation-Parent | ✓ | 46.3% | 3.0% | ✓ | 56.5% | 5.0% |
| Praise | ✓ | 25.4% | 2.8% | ✓ | 28.2% | 3.5% |
| Response Cost | ✓ | 24.9% | 2.8% | ✓ | 23.3% | 2.7% |
| Monitoring | ✓ | 22.8% | 2.8% | ✓ | 24.2% | 0.9% |
| Therapist Praise/Rewards | ✓ | 21.0% | 2.1% | ✓ | 28.5% | 1.9% |
| Family Engagement | ✓ | 19.3% | 1.8% | | | |
| Family Therapy | ✓ | 18.7% | 1.8% | | | |
| Modeling | ✓ | 16.1% | 1.7% | ✓ | 32.3% | 4.3% |
| Goal Setting | ✓ | 40.0% | 1.6% | ✓ | 48.0% | 1.0% |
| Talent or Skill Building | ✓ | 21.4% | 1.6% | ✓ | 21.2% | 1.0% |

| | | | | | | |
|---|---|-------|------|---|-------|------|
| Peer Pairing | ✓ | 22.1% | 1.4% | ✓ | 17.8% | 3.5% |
| Crisis Management | ✓ | 37.0% | 1.3% | | | |
| Educational Support | ✓ | 21.5% | 1.3% | | | |
| Functional Analysis | ✓ | 18.7% | 1.3% | | | |
| Individual Therapy for Caretaker | ✓ | 18.2% | 1.3% | | | |
| Marital Therapy | ✓ | 18.2% | 1.3% | | | |
| Natural and Logical Consequences | ✓ | 19.4% | 1.3% | | | |
| Parent Coping | ✓ | 38.2% | 1.3% | ✓ | 40.3% | 0.7% |
| Activity Scheduling | ✓ | 19.8% | 0.7% | ✓ | 29.4% | 2.1% |
| Self-Monitoring | ✓ | 43.7% | 0.6% | ✓ | 47.6% | 6.9% |
| Behavioral Contracting | ✓ | 22.3% | 0.5% | | | |
| Stimulus Control or Antecedent Management | ✓ | 6.6% | 0.5% | | | |
| Exposure | | | | ✓ | 10.2% | 3.5% |
| Self-Reward/Self-Praise | | | | ✓ | 36.5% | 1.1% |
| Time Out | | | | ✓ | 13.3% | 0.9% |
| Attending | | | | ✓ | 12.4% | 0.8% |
| Commands | | | | ✓ | 12.5% | 0.8% |
| Differential Reinforcement | | | | ✓ | 14.1% | 0.8% |
| Assertiveness Training | | | | ✓ | 16.4% | 0.7% |
| Biofeedback/Neurofeedback | | | | A | 3.7% | 0.4% |

Note. Two minimal sets were identified for Non-Wrap, and checkmarks indicate treatments found in both sets, whereas letters 'A' and 'B' indicate the two treatments of which only one is needed to complete a minimal set.

REFERENCES

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.
- Bernstein, A., Chorpita, B.F., Daleiden, E. L, Ebesutani, C., & Rosenblatt, A. (2011). *How we define treatment affects coverage and efficiency in coordinating evidence-based practice: A relevance mapping analysis*. Manuscript in preparation.
- Bruns, E. J., Sather, A., & Stambaugh, L. (2008). National trends in implementing wraparound: Results from the state wraparound survey, 2007. 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.
- Bruns, E. J., Walker, J. S., Zabel, M., Matarese, M., Estep, K., Harburger, D., Mosby, M., & Pires, S. A. (2010). Intervening in the Lives of Youth with Complex Behavioral Health Challenges and Their Families: The Role of the Wraparound Process. *American Journal of Community Psychology, 46*, 314–331.
- California Department of Social Services (2008). *Wraparound: Background and history*. Retrieved May 30, 2011 from www.dss.cahwnet.gov/cfsweb/PG1320.htm
- Carney, M. M., & Buttell, F. (2003). Reducing juvenile recidivism: Evaluating the wraparound services model. *Research on Social Work Practices, 13*, 551-568.
- Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology, 77*(3), 566-579.

- Chorpita, B. F., Bernstein, A., & Daleiden, E. L. (in press - a). *Empirically Guided Coordination of Multiple Evidence-Based Treatments: An Illustration of Relevance Mapping in Children's Mental Health Services. Journal of Consulting and Clinical Psychology.*
- Chorpita, B. F., Daleiden E. L., Ebesutani C. E., Young, J., Becker, K. D., et al. (in press - b). Evidence-based treatments for children and adolescents: An updated review of efficacy and clinical utility. *Clinical Psychology: Science and Practice.*
- Chorpita, B. F., Daleiden, E., & Weisz, J. R. (2005). Identifying and selecting the common elements of evidence-based interventions: A distillation and matching model. *Mental Health Services Research, 7*, 5-20.
- Clark, H. B., Lee, B., Prange, M. E., & McDonald, B. A. (1996). Children lost within the foster care system: Can wraparound service strategies improve placement outcomes? *Journal of Child and Family Studies, 5*, 39-54.
- Evans, M. E., Armstrong, M. I., & Kuppinger, A. D. (1996). Family-centered intensive care management: A step toward understanding individualized care. *Journal of Child and Family Studies, 5*, 55-65.
- Farmer, E. M. Z., Dorsey, S., & Mustillo, S. A. (2004). Intensive home and community interventions. *Child and Adolescent Psychiatric Clinics of North America, 13*, 857–884.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.
- Glisson, C., Gibbons, R., Green, E. P., Hoagwood, K., Jensen, P. S., Miranda, J., . . . Chorpita, B. F. (2008). Assessing the organizational social context (OSC) of mental health services:

Implications for research and practice. *Administration and Policy in Mental Health and Mental Health Services Research*, 35(1-2), 98-113.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, 9, 329-343.

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic treatment for antisocial behavior in children and adolescents*. New York: Guilford Press.

Hoagwood, K. E., Burns, B. J., Kiser, L., Ringeisen, H., & Schoenwald, S. K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services*, 52, 1179-1189.

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science*, 6(1), 21-37.

Kolko, D. J., Dorn, L. D., Bukstein, O. G., Pardini, D., Holden, E. A., & Hart, J. (2009).

Community vs. clinic-based modular treatment of children with early-onset ODD or CD: A clinical trial with 3-year follow-up. *Journal of Abnormal Child Psychology*, 37, 591-609.

National Advisory Mental Health Council Workgroup on Child and Adolescent Mental Health Intervention Development and Deployment (2001). *Blueprint for change: Research on child and adolescent mental health*. Washington, DC: Author.

New Freedom Commission on Mental Health (2003). *Achieving the promise: Transforming mental health care in America: Final report*. DHHS Pub. No. SMA-03-3832. Rockville, MD: Author

- Rosenblatt, A. (1996). Bows and ribbons, tape and twine: Wrapping the wraparound process for children with multi-system needs. *Journal of Child and Family Studies*, 5(1), 101-116.
- 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(4), 336-351.
- Swendeman, D., & Rotheram-Borus, M. J. (2010). Innovation in sexually transmitted disease and HIV prevention: Internet and mobile phone delivery vehicles for global diffusion. *Current Opinion in Psychiatry*, 23(2), 139-144.
- Task Force on Promotion and Dissemination of Psychological Procedures, Division of Clinical Psychology, American Psychological Association. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, 48, 3-23.
- Tolan, P. H., & Dodge, K. A. (2005). Children's mental health as a primary care and concern: A system for comprehensive support and service. *American Psychologist*, 60, 601-614.
- VanDenBerg, J. E., & Grealish, E. M. (1996). Individualized services and supports through the wraparound process: Philosophy and procedures. *Journal of Child and Family Studies*, 5, 7-21.
- Walker, J. S., & Bruns, E. J. (2006). Building on practice-based evidence: Using expert perspectives to define the wraparound process. *Psychiatric Services*, 57, 127-154.
- Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2006). A proposal to unite two different worlds of children's mental health. *American Psychologist*, 61, 644-645.