

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

Bruce F. Chorpita and Adam Bernstein  
University of California, Los Angeles

Eric L. Daleiden  
PracticeWise, LLC

**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:** The authors demonstrate the feasibility of relevance mapping using data from 1,781 youths in a statewide mental health system and a study data set 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 data set 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 noncoverable 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.

**Keywords:** relevance mapping, coordination, evidence-based treatments

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 EBT 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 one third of all health practice being inconsistent with scientific findings, and more than 20% either unnecessary or harmful (e.g., Agency for Healthcare 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

---

Bruce F. Chorpita and Adam Bernstein, Department of Psychology, University of California, Los Angeles; Eric L. Daleiden, PracticeWise, LLC, Satellite Beach, Florida.

Bruce F. Chorpita and Eric L. Daleiden are members of PracticeWise, LLC, a partnership that provides information resources, training, and consultation services related to evidence-based practices for youth mental health, including the literature database applied in this investigation. Adam Bernstein is a paid consultant with PracticeWise, LLC.

Correspondence concerning this article should be addressed to Bruce F. Chorpita, Department of Psychology, University of California, Los Angeles, P.O. Box 951563, Los Angeles, CA 90095-1563. E-mail: chorpita@ucla.edu

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; Shafraan 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). Those working in mental health services 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—in which a structured comparison of clients from a service population is used 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 analytic effort might not be cost-effective on such a small scale).

### A Missing Link in the “Flow of Knowledge”

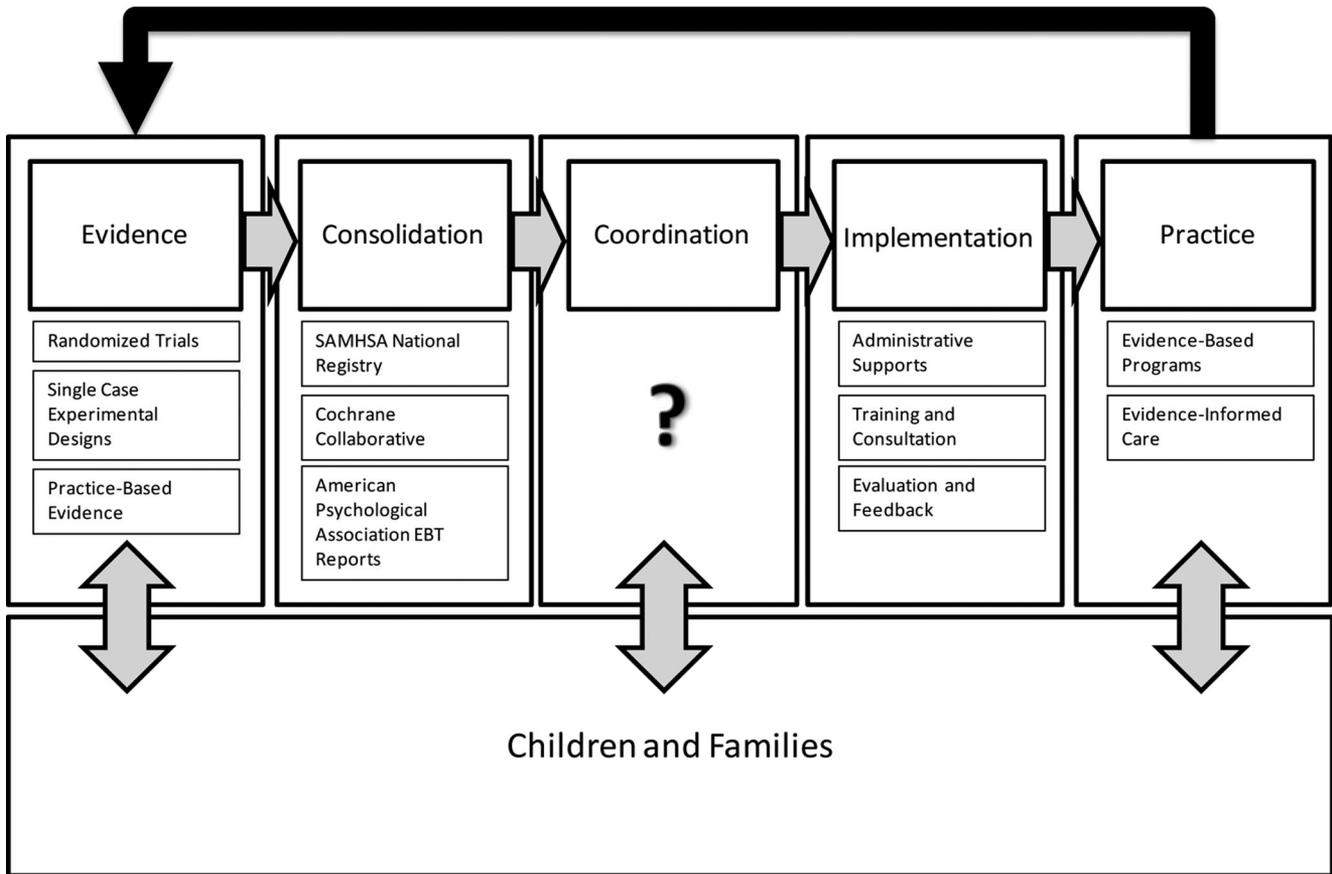
The gap addressed by relevance mapping is illustrated in Figure 1, in which we have positioned a question mark at the juncture at which the flow of knowledge (i.e., the connection between evidence and practice; Frenk, 2009) can break down. The leftmost side of the figure begins with a representation of “raw” evidence, which Graham et al. (2006) 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 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



*Figure 1.* An illustration of the relation of evidence and practice in the context of children's mental health services. Examples are listed for steps for which there are well-defined literatures, products, or concepts. SAMHSA = Substance Abuse and Mental Health Services Administration; EBT = evidenced-based treatment.

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, and 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 data set) 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 data set), 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 data set). 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 data set, and the matching studies are removed from the study data set. 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 was to illustrate relevance mapping as a structured analytic strategy to help guide coordination of treatments, making use of client, study, and workforce data sets 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 boys (64.6%) and 630 were girls (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*, fourth edition, text revision (*DSM-IV-TR*; 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 eight broad categories,<sup>1</sup> 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 eight 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 data sets.** The study data set 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 data set have been reported in previous research (Chorpita et al., in press). Each study had numerous records in the data set, 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 present analysis (e.g., treatment format, therapist education level, etc.). Eight hundred 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 data set 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, a standard of evidence was used, one that has been used over several years with this same service population, which is based largely on the criteria developed and used 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 EBT 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.

<sup>1</sup> The mapping of all diagnoses to problem groups are available upon request.

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	
	$N$	% NC	$N$	% NC	$N$	% 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		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.  $N$ s refer to the total youths within the row label class; percentages refer to the percentage of those youths not coverable. 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.

The problem areas covered by each randomized clinical trial in the study data set were reduced to the same eight broad categories used with the client data set (shown in Table 1). This small number of broad problem categories was chosen after considering alternate problem definitions including specific *DSM-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 trade-off 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 eight 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 trade-off. That said, the 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 data set 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, then those clients are removed and considered coverable prior to the formal relevance mapping analysis. The workforce data set 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 data set (described above) contained 15 studies of MST, and together these studies covered youths of both genders, ages 10–17, 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 data set, 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 data set 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 (see Figure 2, the Practice Units panel). A user may wish to conceptualize treatments as clusters of theoretically related manuals (e.g., family therapy, cognitive behavioral therapy [CBT], 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 “Cop-

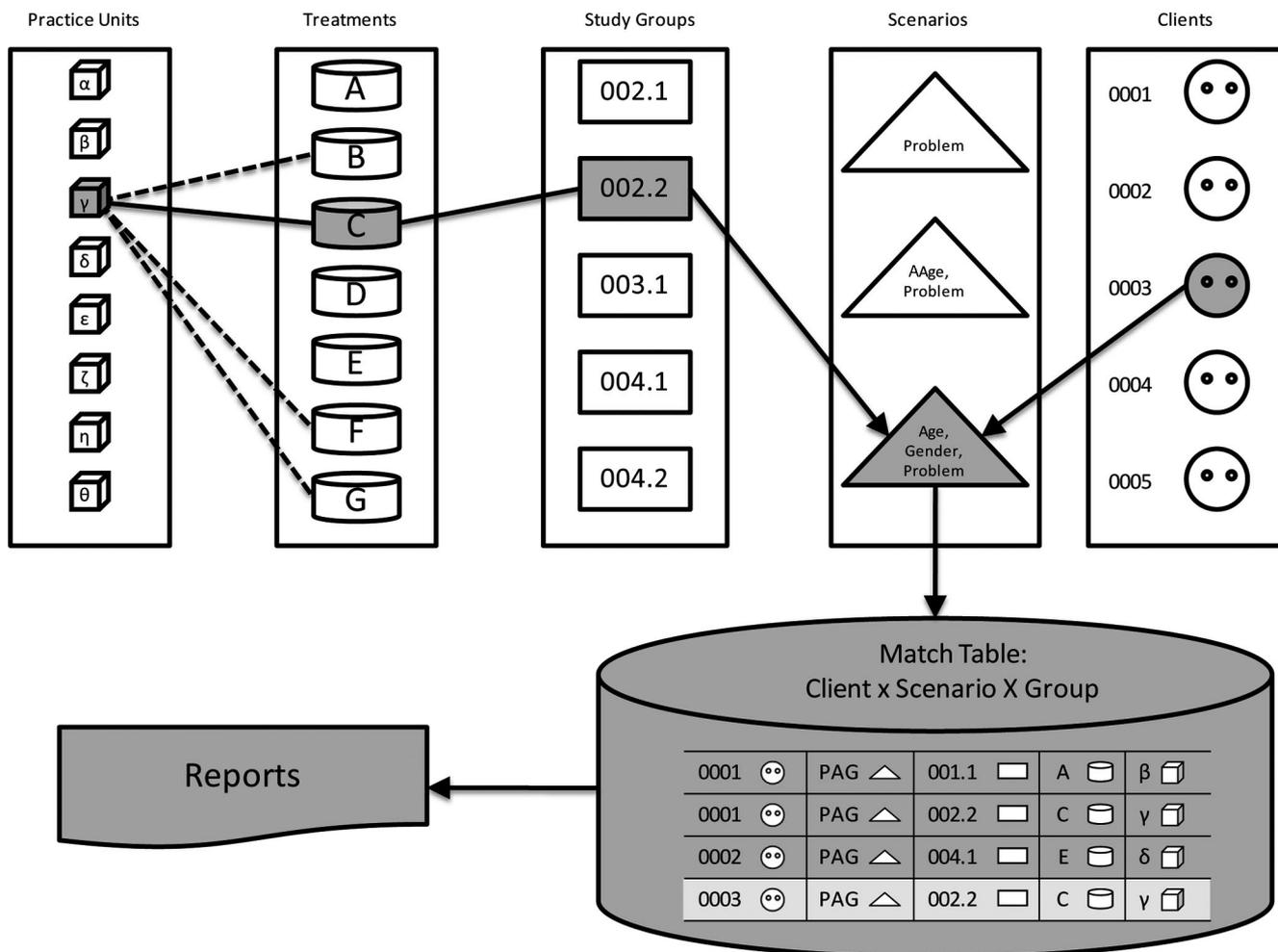


Figure 2. Graphical illustration of the relevance mapping algorithm for determining client coverage. Shaded elements in the top of the figure show the testing of a Single Client × Group × Scenario combination, recorded in the last row of the match table below. In the step of the analysis illustrated, 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  $\gamma$  (e.g., Treatment C could be a specific cognitive behavioral therapy [CBT] protocol, and  $\gamma$  could represent the class of all CBT protocols). PAG = Problem-Age-Gender.

ing 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” vs. matching to “CBT for Anxiety”). Again, relevance mapping does not depend on 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 the present article. As noted above, 832 treatments were grouped 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 data set and the client data set, 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 with 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), then that study group by client combination is written as a record to a “match” data set, 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 (a) the same primary problem as the primary problem of the study participants, (b) any problem (from among several) in common with any problems of the study participants, or (c) 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 data set 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 data set for a given scenario, that is, for whom in the service sample is there *at least one* 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. Noncoverable 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 noncoverable clients). These *client reports* are described as those that refer to clients as dependent variables (e.g., percentage 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.

*Practice reports* are defined in the present analysis 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 with 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, matching for the Problem factor was configured 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, its presence was accounted for 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, CBT, etc.), comparable to clusters used in major lists of effective treatments (e.g., APA Presidential Task Force on Evidence-Based Practice, 2006; Society of Clinical Child and Adolescent Psychology, 2010).

## Results

Table 1 represents a sample client report examining the Hawaii CAMHD data set. 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 data set 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 noncoverable youths with attention problems or autism spectrum problems were noted, presumably owing to their age, gender, or combination thereof (reasons for noncoverability 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 data set (MST) generalized to fewer of the full 1,781 youths under more demanding requirements. Although the 15 studies of MST in the study data set 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 data set in the “Problem Age Gender” scenario (see Table 2) shows that it would be possible to serve the full 58% of “coverable” youths using eight treatment types (plus MST covering the additional 555 youths removed from the residual sample). The first column shows that 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 by any EBT if the treatment in the corresponding row were to be dropped from the set of treatments indicated in that column. 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-deficit/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 0–13. 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 nonoverlapping age ranges (ages 1–10 for ICT and ages 2–12 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 data set that met the standard of evidence for attention-deficit/hyperactivity problems with similar age ranges, though only social skills and medication had demonstrated efficacy for boys, 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%. Because 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 data set, the CAMHD system faced choosing from among an almost incalculable number of possible sets of treatments (e.g., even knowing in advance that exactly eight 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-deficit/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-deficit/hyperactivity, and thus it is the age requirement causing this decrease in coverage). In line with those observations, the study data set did not contain any evidence supporting psychosocial or combined treatments for attention problems for youths above age 13. The uncoverable youths with attention-deficit/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

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.

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 youths for which 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 in which EBTs are less often tested. Notably, even under these strict assumptions, MST was still relevant to more than one 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 noncovered 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 in which 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. IBT, 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 ICT. 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 less 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 researchers' understanding by making visible where inferential leaps in generalizing treatments are large or small. Similarly, relevance mapping makes salient the situations in which 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 in which an otherwise matching EBT has been shown to work. This methodology is extremely flexible in that it is not dependent on 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?). Furthermore, 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 (a) assumptions about the importance of generalization parameters and (b) 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 present 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. Presently, 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 present 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, then 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 on these aggregated parameters are inherently biased to overidentify matches. Thus, in this illustration, estimates of percentage coverable within scenarios 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 data set 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 present 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 data set. A third domain in which 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 Figure 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 data set, 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 mappinglike 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 data set ( $N = 1,781$  youths) from a large state-run service organization, but any coded data set 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.

## References

- Agency for Healthcare Research and Quality. (2001). *Translating research into practice (TRIP)-II*. Rockville, MD: Author. Available from <http://www.ahrq.gov/research/trip2fac.htm>
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.
- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, 61, 271–285. doi:10.1037/0003-066X.61.4.271
- Barlow, D. H. (1981). On the relation of clinical research to clinical

- practice: Current issues, new directions. *Journal of Consulting and Clinical Psychology*, 49, 147–155. doi:10.1037/0022-006X.49.2.147
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66, 7–18. doi:10.1037/0022-006X.66.1.7
- Chorpita, B. F., Daleiden E. L., Ebesutani C. E., Young, J., Becker, K. D., Nakamura, B. J., . . . Starace, N. (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, NY: Springer-Verlag.
- 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: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.
- Frenk, J. (2009, March). *Knowledge translation for health: The marriage of relevance and excellence*. Plenary address at the Third Annual NIH Conference on the Science of Dissemination and Implementation: Methods and measurement. Retrieved from [http://conferences.thehillgroup.com/obssr/DI2010/documents/Plenary%20Session/1\\_Frenk\\_NIH Presentation.pdf](http://conferences.thehillgroup.com/obssr/DI2010/documents/Plenary%20Session/1_Frenk_NIH Presentation.pdf)
- Glass, C. R., & Arnkoff, D. B. (1996). Psychotherapy integration and empirically validated treatments: Introduction to the special series. *Journal of Psychotherapy Integration*, 6, 183–189.
- Glisson, C., Landsverk, J., Schoenwald, S., Kelleher, K., Hoagwood, K., Mayberg, S., . . . The Research Network on Youth Mental Health. (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. doi:10.1007/s10488-007-0148-5
- 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, 13–24. doi:10.1002/chp.47
- 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, NY: 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. doi:10.1176/appi.ps.54.11.1467
- Hromkovič, J. (2004). *Algorithmic adventures: From knowledge to magic*. New York, NY: 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. doi:10.1111/j.1468-2850.1996.tb00072.x
- Kazdin, A. E. (Ed.). (1998). Treatment manuals in clinical practice [Special section]. *Clinical Psychology: Science and Practice*, 5, 361–407. doi:10.1111/j.1468-2850.1998.tb00155.x
- Kendall, P. C. (Ed.). (1999). Clinical significance [Special section]. *Journal of Consulting and Clinical Psychology*, 67, 283–339. doi:10.1037/0022-006X.67.3.283
- Kendall, P. C., & Chambless, D. L. (Eds.). (1998). Empirically supported psychological therapies [Special section]. *Journal of Consulting and Clinical Psychology*, 66, 3–167. doi:10.1037/0022-006X.66.1.3
- Nathan, P. E., & Gorman, J. M. (Eds.). (1998). *A guide to treatments that work*. New York, NY: 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. doi:10.1037/0033-2909.126.6.964
- 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 & Adolescent Psychology*, 37, 8–38.
- Schuster, M. A., McGlynn, E. A., & Brook, R. H. (2005). How good is the quality of health care in the United States? *The Milbank Quarterly*, 76, 517–563. doi:10.1111/j.1468-0009.2005.00403.x
- Shafraan, 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, 902–909. doi:10.1016/j.brat.2009.07.003
- 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 & Adolescent Psychology*, 37, 1–7.
- Society of Clinical Child and Adolescent Psychology. (2010). *Evidence-based mental health treatment for children and adolescents*. Retrieved from <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*. Retrieved from <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(10). doi:10.1037/0003-066X.51.10.1005
- 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. doi:10.1016/j.chc.2004.05.006
- Weisz, J. R., Hawley, K. M., Pirkonis, P. A., Woody, S. R., & Fолlette, 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. doi:10.1093/clipsy/7.3.243
- 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 & Adolescent Psychiatry*, 44, 130–144. doi:10.1097/00004583-200502000-00005

Received September 2, 2010

Revision received April 4, 2011

Accepted April 8, 2011 ■