CrossMark

ORIGINAL ARTICLE

Modeling the Impact of School-Based Universal Depression Screening on Additional Service Capacity Needs: A System Dynamics Approach

Aaron R. Lyon · Melissa A. Maras ·
Christina M. Pate · Takeru Igusa · Ann Vander Stoep

Abstract Although it is widely known that the occurrence of depression increases over the course of adolescence, symptoms of mood disorders frequently go undetected. While schools are viable settings for conducting universal screening to systematically identify students in need of services for common health conditions, particularly those that adversely affect school performance, few school districts routinely screen their students for depression. Among the most commonly referenced barriers are concerns that the number of students identified may exceed schools' service delivery capacities, but few studies have evaluated this concern systematically. System dynamics (SD) modeling may prove a useful approach for answering questions of this sort. The goal of the current paper is therefore to demonstrate how SD modeling can be applied to inform implementation decisions in communities. In our demonstration, we used SD modeling to estimate the additional service demand generated by universal depression screening in a typical high school. We then simulated the effects of implementing "compensatory approaches" designed to address anticipated increases in service need through (1) the allocation of additional staff time and (2) improvements in the effectiveness of mental health interventions. Results support the ability of screening to facilitate more rapid entry into services and suggest that improving the effectiveness of mental health services for students with depression via the implementation of an evidence-based treatment protocol may have a limited impact on overall recovery rates and service availability. In our example, the SD approach proved useful in informing systems' decision-making about the adoption of a new school mental health service.

Keywords Screening · Depression · Implementation · System dynamics modeling · School mental health

A. R. Lyon (✉) · A. Vander Stoep
Department of Psychiatry and Behavioral Sciences, University of Washington, 6200 NE 74th St., Suite 100, Seattle, WA 98115, USA
e-mail: lyona@uw.edu

M. A. Maras
University of Missouri, Columbia, USA

C. M. Pate
Health & Human Development Program, WestEd, Los Alamitos, USA

T. Igusa
Johns Hopkins University, Baltimore, USA

Introduction

There is growing support for the utility of universal mental health screening programs in middle and high schools to identify youth who could benefit from—but may not ordinarily receive—emotional and behavioral health services (McCormick et al. 2009; U.S. Public Health Service 2000; U.S. Department of Education Office of Special Education and Rehabilitative Services [OSEP] 2002). Research suggests that while nearly 1 in five school-age youth in the United States have a diagnosable psychiatric disorder (Merikangas et al. 2010; Shaffer et al. 1996), the majority of mental health problems are undetected and untreated (Kataoka et al. 2002). Internalizing symptoms of depression or anxiety are particularly likely to go unnoticed by school personnel, parents and other adults (Bradshaw et al. 2008; De Leo and Heller 2004). As a result, students with internalizing difficulties are significantly under-referred by teachers (Walker et al. 2005, 2010) and underserved by special

education and mental health systems (Bradshaw et al. 2008). When internalizing symptoms persist untreated, they typically intensify, resulting in more severe mental health problems (e.g., psychiatric disorders, suicide), as well as academic and other functional impairments (Best et al. 2004; Fletcher 2008; Substance Abuse and Mental Health Services Administration [SAMHSA] 2008; Wilcox-Gok et al. 2004). Universal depression screening followed by evaluation and referral of positive screens can ameliorate the under-detection and lack of attention to this prevalent and often debilitating condition.

Traditional approaches to the identification of student problems (e.g., office discipline referrals) tend to yield non-compliant and disruptive behaviors while underemphasizing internalizing concerns (Walker et al. 2005, 2010). In contrast, universal screening for depression symptoms has been demonstrated to increase identification rates for indicated service provision (Eklund and Dowdy 2014; Scott et al. 2009; Williams et al. 2009). When paired with timely interventions, early identification of depression has the potential to improve emotional and functional outcomes for youth. Accordingly, universal screening is supported by numerous professional organizations [e.g., American Educational Research Association (AERA), American Psychological Association (APA)], meets guidelines and mandates set forth by government agencies (e.g., Surgeon General's Report, U.S. Department of Health and Human Services 1999; New Freedom Commission on Mental Health 2003), is generally cost-effective (Kuo et al. 2009) and supports laws established by the federal government [e.g., Individuals with Disabilities Education Act (IDEA) of 2004].

Universal depression screening is typically implemented in a series of steps: Optimally, parents, students, teachers, and school staff are informed with regard to the rationale for universal depression screening and the screening procedures. Parents and students voluntarily elect to participate, after which screeners administer a validated questionnaire to students in classroom settings. Questionnaires are scored and students who "screen positive" for depression undergo follow-up assessment. Some students are determined to be distressed and in need of mental health intervention. These students are offered support on site or referral to community services. Parents are engaged to help students execute the follow-up support plan.

Barriers to Large-Scale Screening Implementation

Despite their potential value, studies indicate that only 15 % of schools conduct any sort of school-wide health screening (Foster et al. 2005) and only 2 % carry out universal screening for emotional and behavioral problems (Romer and McIntosh 2005). A variety of barriers, both perceived and actual, impede efforts at widespread implementation of universal mental health screening in schools. These include concerns about the usability and practicality of screening tools, including developmental, cultural, and contextual appropriateness of specific measures; general concerns about reliability and validity of screening measures and implications regarding false positives and negatives (Dever et al. 2012; Fox et al. 2008); legal and ethical issues related to parental consent and confidentiality; and fear of stigmatizing students. Prominent among the commonly-referenced barriers are concerns that the students identified may exceed schools' service delivery capacities, placing educational institutions in the undesirable situation of generating critical, "actionable" information about student distress and having inadequate ability to respond (Severson et al. 2007; Weist et al. 2007). Although research has addressed many of the concerns about the potential negative consequences of introducing screening programs listed above (e.g., in support of the cross-cultural appropriateness of some established measures; Banh et al. 2012; means of minimizing stigmatization; Vander Stoep et al. 2005), the service capacity concern has persisted largely untested.

The issue of service capacity translates to the question of who will provide support to address the hidden needs that may be systematically uncovered by screening. School personnel—including counselors, nurses, teachers, and administrators—play an integral role in promoting positive student behavior and social and emotional development, yet in most schools, these personnel are not trained nor mandated to provide mental health services. For school personnel who are trained, time to devote to student mental health issues is typically insufficient. For instance, national guidelines recommend a ratio of one school counselor to 250 students; however, data indicate the national average is nearly double that [American School Counselor Association (ASCA) 2011]. Furthermore, Foster and colleagues (2005) documented that the modal number of mental health personnel per school nationwide is three and that, on average, these professionals (including school counselors, school psychologists, school social workers, school mental health counselors, school nurses, etc.) spend no more than half of their time providing mental health services. In light of these findings, concerns that the number of students identified may exceed schools' capacities to provide mental health interventions are warranted and can translate into reluctance to implement universal depression screening (Severson et al. 2007). Addressing concerns about increased service need is necessary to support the adoption of universal screening in schools.

Although service delivery capacity is a perceived barrier to implementing universal mental health screening in schools (Severson et al. 2007; Weist et al. 2007), no research has systematically evaluated this perception. Researchers encounter difficulties when attempting to quantify the

impact of novel programs within complex service systems, where they interact with existing programs, policies, and norms at a variety of levels (Aarons et al. 2011; Damschroder et al. 2009). Indeed, the introduction of universal depression screening into schools is likely to have an array of proximal and distal service implications, the nature and scope of which may not be immediately apparent. For instance, even as a greater number of youth are identified as needing intervention, the extent to which this otherwise unidentified group will result in long-term, noticeable strain on a school's service capacity may be mitigated by factors such as the likelihood that services can effectively and efficiently address their needs. Careful consideration of these variables and their interrelationships will be necessary to determine the actual impact of depression screening and to identify leverage points, which may assist stakeholders in understanding how even small shifts in one part of a complex system can produce larger changes in other parts of the system (Forrester 2009; Meadows 1999).

Addressing Increased Service Need

Sentiment that schools provide a natural setting for mental health prevention and promotion is growing (Adelman and Taylor 2007; Weist 2008), but no studies have evaluated the potential utility of specific service improvement approaches that could compensate for anticipated increases in service need resulting from universal depression screening. Below, we describe two potential compensatory approaches: the *service capacity approach*, in which resources are invested in increasing the number of providers delivering mental health care, and the *evidence-based practice (EBP) implementation approach*, in which resources are invested in improving the quality of care delivered by existing providers.

Service Capacity Approach

Although some promising approaches are being developed that use new technologies to augment the care delivered by a relatively small number of mental health providers (e.g., computer-assisted psychotherapies; Becker and Jensen-Doss 2013), most efforts to enhance service *quantity/availability* involve the introduction of additional service providers into an existing system, often at substantial cost. In the context of universal depression screening, increasing the quantity of services available in a school to compensate for the potential influx of students identified could involve hiring one or more part- or full-time mental health providers to deliver services to students who screen positive. Nevertheless, the actual impact of increasing the number of service providers on service utilization is unsubstantiated. Notably, Green and colleagues (2013) recently found that

the ratio of students to mental health providers in a school was not associated with overall service use across sectors. Furthermore, screening positive for depression does not always translate into a need for specialty mental health services. Systematic follow-up of students who screen positive to determine the source of their distress has shown more needs for academic (e.g., tutoring, homework help) and social (e.g., after school activities) support than for specialty mental health services (Vander Stoep et al. 2005), underscoring the importance of considering the need for expansion of non-mental health services when anticipating the impact of universal depression screening.

EBP Implementation Approach

Researchers and administrators have also begun advocating for strategies that improve service *quality* to better support the mental health needs of entire populations. The use of high-quality, research supported interventions is expected to enhance efficiency and limit costs by increasing the likelihood and rate of client improvement, thus reducing the average length of episodes of care and recurrent needs for care. A variety of well-researched, evidence-based treatments have been shown to outperform the "usual care" mental health services commonly delivered to youth by providers working in community settings (Weisz et al. 2006a). For instance, multiple effective treatment approaches for adolescent depression have been identified (Weisz et al. 2006b). In conjunction with universal screening, these approaches could significantly enhance the quality of services, resulting in more efficient and effective care (Severson et al. 2007).

In the context of universal depression screening, increasing service *quantity* (i.e. the service capacity approach) and service *quality* (i.e., the EBP implementation approach) are both potentially viable options, worthy of further study. To date, no published articles describe comparisons of the potential utility of these alternative approaches to addressing additional service needs generated through universal depression screening. Furthermore, the complex interplay among system variables makes it difficult to evaluate the effectiveness of either compensatory approach, necessitating new analytic methods that quantitatively examine complex and recursive interactions among service system components over time.

System Dynamics Modeling

Implementation of universal depression screening in a school setting is likely to create ripple effects throughout multiple ecological systems (classroom, family, school, community). Thus, in considering the implications of identifying student needs that are otherwise undetected, a systems or ecological view is warranted. System dynamics (SD)

is a methodological approach rooted in systems science (Mabry et al. 2008; Sterman 2000), with applicability to the implementation of new interventions in community contexts (Hovmand 2013). The approach can be used to make implicit assumptions explicit, identify knowledge gaps, and locate leverage points for system change (Mabry et al. 2008; Meadows 1999). In the context of implementation, SD can be used to simulate the outcomes of introducing new programs under different conditions/assumptions, thus providing some degree of empirical evidence to support or refute untested or theoretical approaches (Levy et al. 2010; Sterman 2006). Instead of reducing models to simple linear pathways, a SD modeling approach allows for the mapping of recursive influences within and between systems as well as over time (Hirsch et al. 2007). Feedback loops may intensify, lessen, or otherwise alter the relationships among different processes within a system, generating novel patterns that are greater than the sum of their parts (Forrester 2009). Although such feedback loops are frequently referenced within implementation models (e.g., Aarons et al. 2011; Damschroder et al. 2009), there are few approaches available to model them quantitatively (Warren 2004). For example, these models often describe recursive information flows between implementation stages as knowledge about maximizing the compatibility of a program in a destination context is gained in later stages and fed back into earlier-stage processes, thus enhancing fit (Aarons et al. 2011; Fixsen et al. 2005).

At its most basic level, SD modeling involves the specification of three components (Forrester 2009): *stocks* (variables that increase or decrease over time), *flows* (rates of change in a stock), and *feedback loops* (connectors of stocks and flows over time). Stocks are the primary outcome of SD models and can represent any variable that is expected to change as a result of the specified flow and feedback loop parameters. For instance, stocks may be informational (e.g., service provider knowledge or skill) or represent quantities of tangible resources (e.g., number of service recipients receiving an intervention; amount of professional time devoted to a task). All stock–flow systems share the same underlying structure—the resource level (stock) is modified based on the movement of inflows to it and outflows from it (Forrester 2009). Using a water analogy, the stock is a storage tank, with an input pipe above and a drain below, each controlled by valves. The water inflows and outflows change the stock. If the drain and input valves were to close, there would be no change in the water level. However, if the drain valve were to open just slightly, then the tank (stock) would gradually deplete. Simultaneously, if the input valve is opened further than the drain valve, then the tank level would increase. If maintaining water in the tank is desirable, input and output valves can be opened an equal amount.

SD modeling has recently been applied to health services research and implementation science (Homer and Hirsch 2006; Leischow and Milstein 2006). The approach allows stakeholders to ethically and cost-effectively explore hypothetical scenarios to guide decision-making and implementation in a manner that is tailored to the local context and responsive to the needs of stakeholders (Tobias et al. 2010). Complementing other methods, SD has the potential to translate system complexity into simple, practical actions and to predict intended and unintended consequences and time-delayed effects (Mabry et al. 2008).

Aims/Purpose of the Study

The overarching goals of the current study are twofold: (1) to provide a detailed example of the utility of SD modeling (Hirsch et al. 2007) to inform decisions about implementing new practices in community-based service settings and, in doing so, (2) explore how SD modeling may be used where research is lacking to inform decisions about implementing school-based depression screening. Drawing existing data from the youth mental health services literature, our study uses a SD modeling approach to simulate the impact of a universal depression screening program in a high school setting with the goal of identifying (a) system components that may influence service delivery and outcomes over time, (b) additional resources that may be needed to facilitate such a program or mitigate any negative effects, and (c) key leverage points that provide efficient opportunities for addressing youth needs. In doing so, the following research questions will be addressed:

(1) What is the anticipated impact on service need and service use of introducing a universal depression screening program into an "average" high school (based on size, prevalence/severity of student mental health problems, etc.)?

(2) What are the anticipated effects of two different *service improvement approaches* (increased service capacity; EBP implementation) intended to compensate for the heightened service demand that may result from universal depression screening?

In addition to these formal research questions, the study will also identify additional data that are needed to more effectively model the impact of universal depression screening on school and larger system resources, as well as the effectiveness of compensatory approaches. Although it was anticipated that the current project would shed light onto some of the most commonly-articulated concerns about universal emotional health screening, it was not intended to be the "final word" on the legitimacy of those concerns or the effectiveness of the compensatory approaches tested. Instead, it is our goal to present a valid

and efficient method for addressing key questions without the need for novel data collection.

Method

Information Sources

This project used information drawn from the existing empirical literature on mental health services, emotional health screening, school mental health, and evidence-based treatments for depression to construct a series of quantitative SD models that address the research questions listed above. A non-exhaustive literature search was conducted to identify representative example data that could be used to specify model components. Searches were conducted independently for each model element (e.g., adolescent depression prevalence, school-based mental health service availability, rates of school-based mental health service utilization), but the search was discontinued when appropriate, representative, and plausible data were located. When possible, larger-scale studies, meta-analyses, and national surveys (e.g., Foster et al. 2005) were prioritized to maximize the relevance and ecological validity of the resulting models. Because all information included was available in the existing scientific literature, and no original data collection was conducted, this study was not subject to institutional review board approval.

To accomplish the goals of the current paper and yield useful information, it was necessary to identify information sources that could fulfill model components relevant to service need, likelihood of service utilization, service capacity, and intervention effectiveness before and after the introduction of universal screening. Information sources (Table 1) were identified to fulfill a number of core model information needs including: (a) prevalence of clinically-significant or borderline significant depression within a student body (13.9 %); percentage of (b1) depressed and (b2) non-depressed students likely to have used mental health services, having been identified through means other than a depression screening program (5.8 and 14.2 %, respectively); (c) rates of universal screening participation (82.5 %); (d1) true and (d2) false positives, based on the identified depression screening tool (see below) (79 and 21 %, respectively); (e) mental health service provider staff capacity [1.0 full-time employment (FTE) in the base model, spending 85 % time on mental health treatment delivery and 15 % on initial assessment sessions]; (f1) likelihood of entry into mental health treatment or (f2) an alternative pathway that included no further intervention or other non-mental health supportive services for students for whom services were indicated (35.8 % mental health, 64.2 % alternative pathway); (g) mental health treatment

duration (12 weeks); (h1) recovery and (h2) remission rates for "services as usual" (52 and 48 %, respectively); and (i) recovery rates for the identified evidence-based treatment (74 %). These information categories and the full list of studies identified to fulfill these needs can be found in Table 1.

Model Building

A graphic representation of the model, shown in Fig. 1, was constructed using Vensim PLE (Ventana Systems 2006). The boxes in the figure represent stocks of students and the double arrows represent channels of flow between the stocks, with flows regulated by "valves," shown as double triangles. The single arrows indicate the parameters that influence the flows. With this representation, the flow of students, levers for flow regulation, and feedback caused by recovery can be visualized. This model is dynamic, in that flows of students vary each week, based on a number of model parameters (e.g., time spent in treatment).

Using the data sources identified in Table 1, we constructed models for a hypothetical public high school with 1,000 enrolled students over the course of one academic year (i.e., 36 weeks). In the current application of SD, individual students populated the stocks, and the various literature-driven system parameters listed above (data elements *a–i*) controlled the system flows and feedback loops. Below, we describe our model-building steps and various versions of the model, with references to the literature sources that informed their construction. The algorithmic details of the model can be found in the Appendix.

Base Model

First, a *base model* was developed to reflect "business as usual" without systematic identification of students in need of mental health treatment. To construct the base model, we used data elements *b* (service use without screening), *e* (staff capacity), *f* (entry into mental health treatment versus an alternative pathway/other services), *g* (treatment duration) and *h* (recovery rates for services as usual). In the base model and all subsequent models, staff capacity (*e*) was allocated to two direct service activities: (1) initial assessments of newly referred students and (2) mental health treatment for students with some level of identified mental health problem. Initial assessment sessions were included to account for findings that the majority of students who use school mental health services do so on a short-term basis, only attending one session (Walker et al. 2010). Services provided in these initial sessions are anticipated to include in-depth assessment of a problem and determination of a diagnosis/primary presenting problem, crisis management, and triage/determinations about next intervention steps or

Table 1 Model components and information sources

Model information component	Identified value	Source
(a) Prevalence of clinically-significant or borderline depression within student body	13.9 %	13.9 % of students may screen high for symptoms of depression (i.e., 20+ on the MFQ; Vander Stoep et al. 2005)
(b) Percentage of student body likely to use mental health services over a one-year period for (i) depression and (ii) other reasons prior to the introduction of a screening program	(b1) 5.8 % (depressed);	*Depressed* 40–43 % (avg: 41.5 %) of students with depression symptoms have recently used MH services in school (Lyon et al. 2013) $13.9\% * 41.5\% = 5.8\%$
	(b2) 14.2 % (non-depressed)	*Non-depressed* 20 % of students access school-based mental health services in a given year (Foster et al. 2005) $20\% - (13.9\% * 41.5\%) = 14.2\%$
(c) Rates of universal screening participation	82.5 %	80–85 % (avg: 82.5 %) screening participation rate may be expected ($\sim$10 % refusal, 10 % no response), using passive consent procedures (Chartier et al. 2008; Kuo et al. 2009)
(d) True and false positives, based on the identified depression screening tool (MFQ)	(d1) 71 % (true positives) (d2) 29 % (false positives)	Identified true and false positive rates using the MFQ for school-based depression screening (Vander Stoep et al. 2005)
(e) Mental health service provider staff capacity	22.5 direct service hours/week	Brener et al. (2007) documented that 77.9 % of schools have at least one part-time counselor who provides mental health services. We included a 1.0 full-time employment (FTE) provider and calculated that this providers would be able to see 20–25 students a week (avg: 22.5), assuming that students are seen weekly for 45–60 min
(f) Likelihood of entry into mental health treatment or other supportive services for "true positive" students receiving an assessment session	(f1) 35.8 %/(f2) 64.2 % (mental health treatment/non-mental health intervention)	Because the primary focus of this study was on staff capacity to deliver indicated mental health treatment and because non-mental health services vary widely (e.g., academic, after-school, mentorship, etc.), non-mental health services were modeled with unlimited capacity
(g) Mental health treatment duration	12 weeks	To enhance the comparability across models, mental health treatment duration was set to 12 weeks, based on the length of IPT-A and usual care intervention in the identified IPT-A studies (Mufson et al. 1999; Rosselló and Bernal 1999)
(h) Recovery/remission rates for services as usual	(h1) 52 %/(h2) 48 %	Evidence suggests that community-treated depressed youth look similar to control youth from RCTs in their change trajectories (Weersing and Weisz 2002). In a depression treatment study, 52 % of youth receiving school-based usual care for depression (control condition) experienced clinically-significant improvement (Mufson et al. 2004a, b)
(i) Recovery rates for the identified evidence-based treatment (IPT-A)	74 %	Results from randomized trials indicate that 74–82 % of youth receiving Interpersonal Therapy for Adolescent Depression (IPT-A) experience clinically-significant improvement (Mufson et al. 1999; Rosselló and Bernal 1999). Given that the effectiveness of interventions often attenuates when implemented outside of the context of research projects, we selected the lower number as our more conservative estimate

referrals to other programs. In our models, all students referred to mental health services (via traditional pathways or through screening) first completed an initial assessment session. In the base model, students were then referred to either mental health treatment or an alternative pathway that could include non-mental health interventions or supports. The alternative (i.e., non-mental health support) pathway was designed to represent both indicated non-mental health interventions (e.g., tutoring, mentoring) as well as school-wide, universal interventions designed to support the mental health and well-being of all students (e.g., mental health awareness campaigns, promoting positive school climate). In this way, our model artificially constrained the diverse range of service use patterns commonly seen in schools to these two pathways (a single, initial assessment pathway or mental health service delivery for 12 weeks). Although the 12-session pathway did not explicitly attend to students who might need between 2 and 11 sessions or those who needed
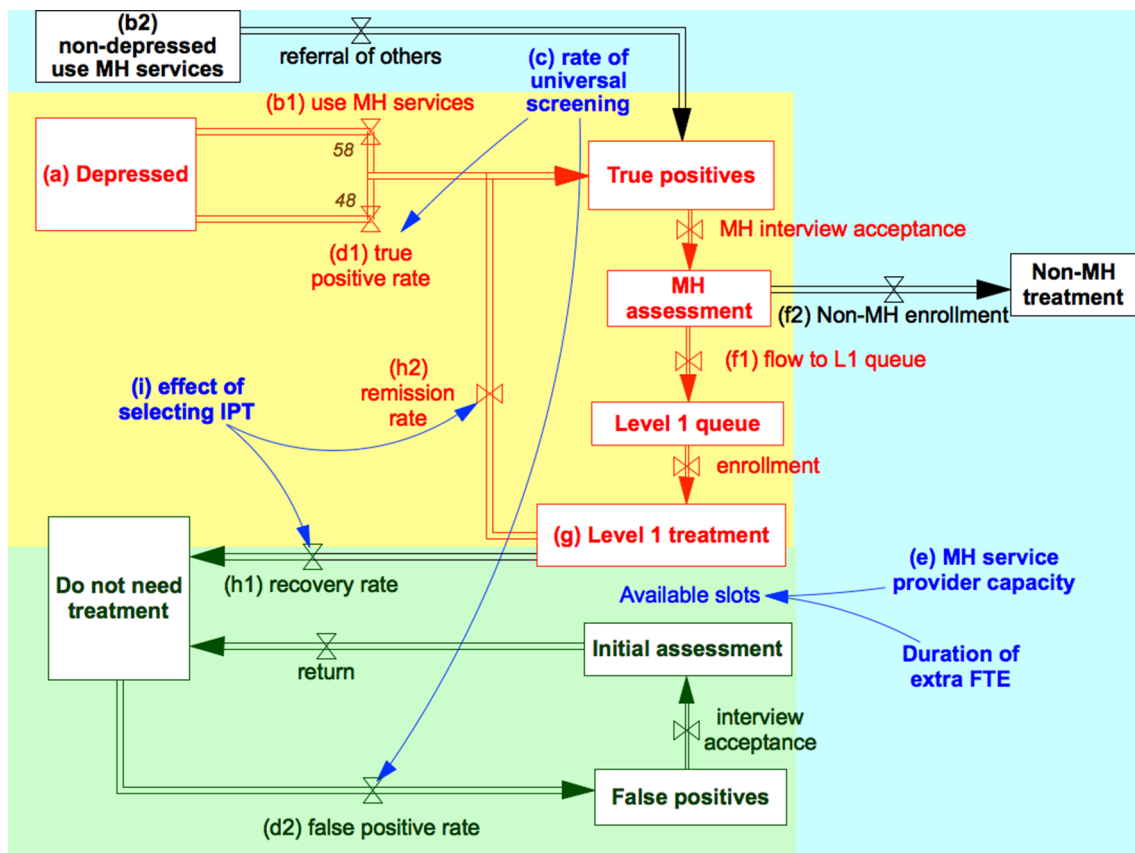
Fig. 1 Full system dynamics screening model

13 or more sessions, it was intended to approximate the average service capacity demand that typically varies between two and over thirty sessions over the course of the academic year (Walker et al. (2010) found that some students may receive over 100 school mental health treatment sessions in a single year). Weekly staff capacity was allocated at 15 % to initial assessment sessions with new referrals and 85 % to mental health treatment for a growing caseload of students whose initial assessments indicated the need for a full episode of treatment. However, when insufficient students were available for indicated mental health treatment (e.g., at the beginning of the year), our models allowed remaining clinical staff capacity to be allocated to additional initial assessment sessions.

Universal Screening Model

Next, the base model was expanded to include systematic identification of students in need of mental health treatment. The *universal screening model* incorporated additional elements *a* (prevalence of depression symptoms), *c* (proportion of enrolled students who participate in screening), *d* (proportion of true and false positives for the identified screening tool), and *f* (likelihood of service

entry). To identify relevant data for element *d*, the Mood and Feelings Questionnaire (Angold et al. 1987) was selected because its reliability and validity have been supported by decades of research, it is available in the public domain—and therefore accessible to professionals working in low-resource settings—and it was the primary screening tool in a subset of the studies used to construct our models (e.g., Banh et al. 2012; Lyon et al. 2013; Vander Stoep et al. 2005). In the universal screening model, an additional pathway was included to account for false positives that screened positive but were not actually experiencing sufficient distress to be considered for services of any type. After an initial assessment session, these youth are returned to the general school population stock.

Universal Screening + Service Improvement Models

The universal screening model was then adjusted to assess the impact of the two service improvement/compensatory approaches identified: (1) the service capacity expansion approach and (2) the EBP implementation approach. Hiring additional mental health clinicians to provide services was expected to address the increased demand for "hands on deck" due to students whose undetected depression was

brought to light by a positive screen. Training the existing mental health clinician to implement an empirically supported treatment was expected to more effectively address the needs of depressed students whose needs were identified through usual means and via universal screening, thereby reducing the prevalence of depression in the student population.

Service Capacity Expansion In the service capacity expansion approach, we modeled the effects of increasing the number of service providers in the school to address the additional need for services hypothesized to result from the introduction of universal screening. Separate models were run for scenarios in which one, two, and three full-time mental health service providers were available (relative to only one service provider in the base model). In these scenarios, the breakdown of service provider activities remained the same as in the base model (i.e., 15 % initial assessments, 85 % long-term service provision). In addition, because it was expected that much of the increased service need resulting from screening would be time-limited to the immediate post-screening period, a series of models were constructed in which new service providers were introduced to the school on a short-term basis following screening to conduct initial assessment sessions. In these scenarios, combinations of different durations of short-term provider availability (e.g., 2, 4, 10 weeks) and different numbers of providers (i.e., one, two, three) were evaluated to determine the most efficient method of addressing the post-screening service need increase.

EBP Implementation Models exploring the EBP implementation approach evaluated an alternative method of addressing the anticipated increase in service need. In this scenario, we tested the impact of the introduction of an established evidence-based intervention to existing school staff for use with depressed youth, Interpersonal Psychotherapy for Adolescent Depression (IPT-A; Mufson et al. 2011). IPT-A was selected due to strong evidence for its effectiveness with adolescents experiencing depressive symptoms (Mufson et al. 1999, 2004b; Rosselló and Bernal 1999), previous applications in schools (e.g., Mufson et al. 2004a; Young et al. 2006), and high ratings for both effectiveness and implementation readiness on the SAMHSA National Registry of Evidence-Based Programs (www.nrepp.samhsa.gov). IPT-A is a 12-week intervention, and efficacy studies suggest that between 74 % (Mufson et al. 2004b) and 82 % (Rosselló and Bernal 1999) of youth receiving IPT-A are likely to "recover" (i.e., move from the clinical to nonclinical range) during treatment. In the current project, we selected the more conservative of these numbers (i.e., 74 %) for model construction to avoid overestimating the impact of the

intervention and because it reflected the results of a study conducted in a school setting. In the EBP implementation model, this value replaced the original recovery rate of 52 % selected for usual care for the subset of youth service recipients anticipated to have depression as their primary presenting problem (data element h).

Visual Depiction of Full Model

Figure 1 displays the full model for the total population of students ($n = 1,000$). In the base model, flows for the universal screening process (i.e., true identification, false identification) were simply "turned off" (i.e., set to zero). Following their construction, each model was run to evaluate a set of longitudinal and cumulative outcomes, which allowed for direct output comparisons across scenarios. Longitudinal outcomes are those for which the pattern of change over time (i.e., over the full academic year) is most meaningful. These included (a) students waiting for an initial assessment session who will ultimately *need mental health intervention*, (b) students waiting for an initial assessment session who ultimately *need non-mental health services*, (c) students who have completed an initial assessment session and are *waiting for mental health intervention*, (d) provider hours per week devoted to *providing mental health intervention*, (e) provider hours per week devoted to initial assessment *meetings with true positives*, and (f) provider hours per week devoted to initial assessment *meetings with false positives*. In contrast to longitudinal outcomes, the values of the cumulative outcomes are most meaningful at the end of the year, when the effects of each model have had the opportunity to compound. Cumulative outcomes all involved the number of students who had populated a given stock at some point during the 36 week academic year. These included number of students (a) receiving mental health treatment, (b) recovering from their mental health problem, and (c) receiving non-mental health services. The details of the algorithm that underlies each of the models described above are provided in Appendix.

Results

Base Model

Results from the pre-screening base model are displayed in Fig. 2. In this model, the school has a single full-time mental health service provider, and students are referred gradually for services over the course of the academic year. As stated earlier, in all the models the time breakdown of practitioner responsibilities was specified to include at least 15 % devoted to initial assessment meetings with all remaining
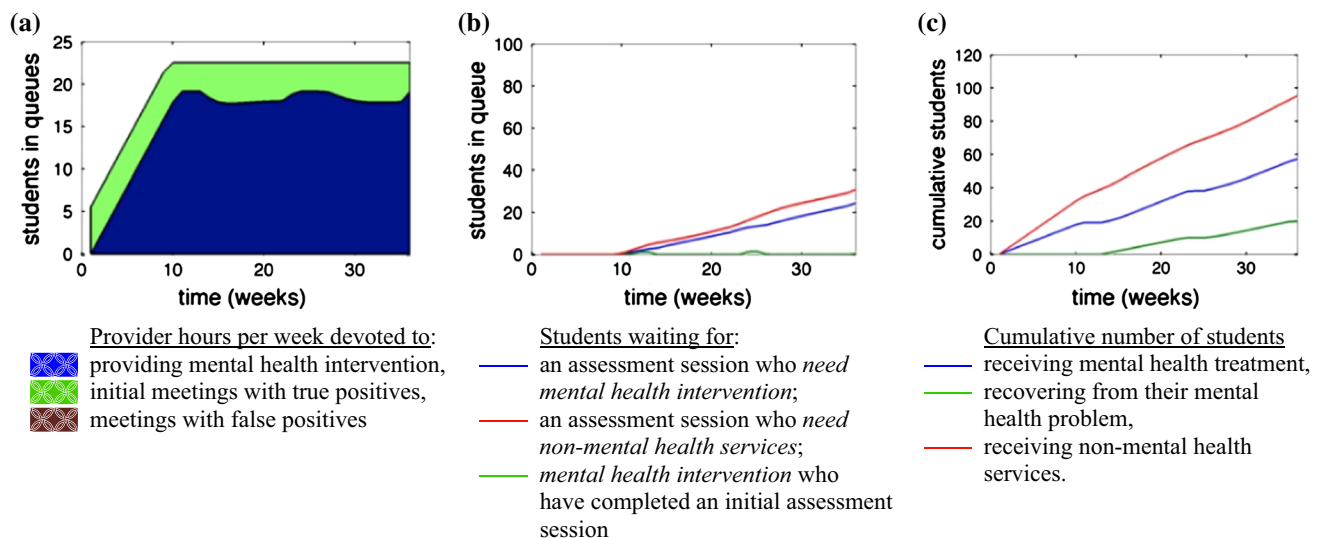
Fig. 2 Effect of no screening

time devoted to mental health treatment. Figure 2a displays an area graph of the breakdown of providers' work responsibilities over time. Because referrals are gradual (approximately 5–6 per week), providers are initially able to accommodate the assessment and treatment needs of referred students without reaching full capacity. With each passing week, however, a subset of the students who have completed an initial assessment session are identified for entry into mental health treatment; indicated by the growing area in the graph representing students receiving treatment. By the tenth week of the school year, the service provider has reached her full capacity. Because the treatment of existing cases is prioritized in the model, she can now only devote 15 % of her time to initial assessment sessions with new cases, and a backlog of referred students begins to grow. Beginning at week 13, a "wave" pattern can be observed as students complete their 12-week intervention sequence, enabling the clinician to briefly devote a higher percentage of her time to initial assessment sessions. Ultimately, these assessment sessions yield a sufficient number of students requiring mental health intervention, again filling all but the reserved 15 % of the provider's caseload by week 24.

As can be seen in Fig. 2b, the backlog begins around week 10 and includes students who will ultimately receive mental health treatment, as well as those who will be referred to non-mental health programs. This backlog will continue to grow for the remainder of the academic year. By the end of the year, 24 students in need of mental health treatment and 31 students destined for the alternative pathway (including non-mental health services) will remain in the initial assessment queue, never receiving an assessment and triage session. However, nearly all students who have completed an initial assessment and been assigned to mental health treatment will be accommodated.

Among those who have completed an initial assessment, Fig. 2c demonstrates that, over the course of the year, 57 total students will receive mental health treatment, 20 of whom can be expected to experience lasting clinically-significant benefits. Furthermore, an additional 95 students will be referred to non-mental health services.

Universal Screening Model

The universal screening model evaluated the impact of school-wide administration of the MFQ on service need and staff capacity. The model was constructed with an assumed screening participation rate of 82.5 % of the student body (based on findings suggesting that participation rates range from 80 to 85 %; Vander Stoep et al. 2005; Chartier et al. 2008; Kuo et al. 2009; Table 1). In the screening model, the effects on staff responsibilities are nearly immediate. As indicated in Fig. 3a, all unused staff capacity in the early weeks of the school year in the base model is now devoted to conducting post-screening initial assessment sessions. Initially, a large portion of the student assessment sessions are devoted to meetings with false positive screens, but these decline relatively quickly. Interestingly, the screening process also appears to slow the rate of treatment entry for some youth, as the provider's mental health treatment caseload does not reach capacity until 13 weeks (3 weeks later than in the base model). After the peak in treatment caseload size in week 13, more initial assessment session capacity is opened, but similar to the base model, this gradually closes as a subset of the students who have received an initial assessment session are accepted into mental health treatment.

In the screening model, establishment of a queue of backlogged youth in need of initial assessment sessions is
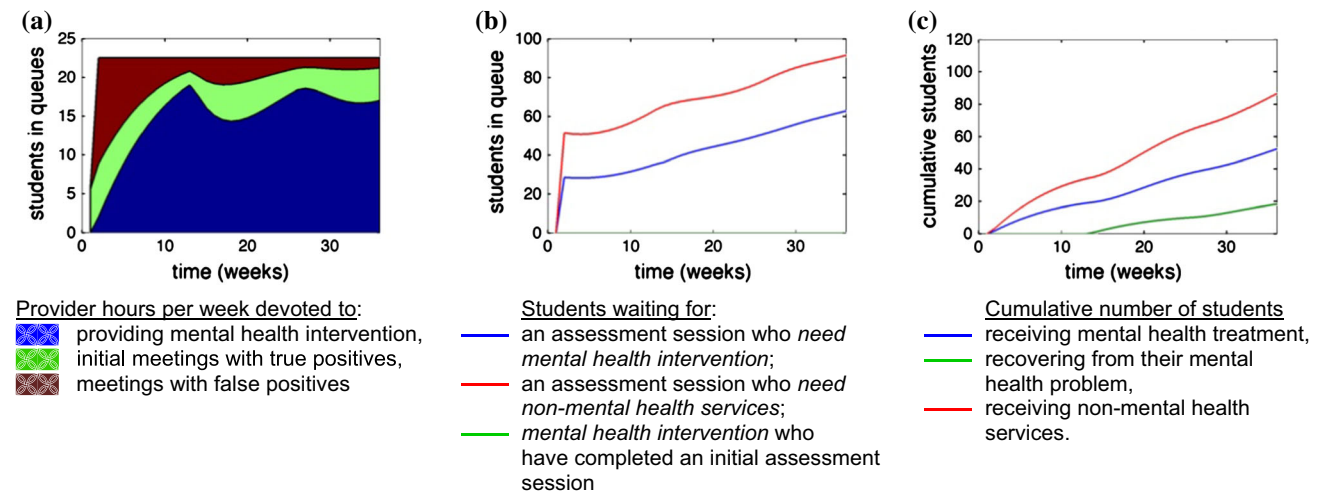
Fig. 3 Effect of universal screening

nearly immediate, with approximately 29 youth needing mental health treatment (and 51 needing non-mental health services) waiting for an assessment session after the first week of school (Fig. 3b). This number also continues to climb over the course of the school year, ending the 36 week period with 63 youth who are in need of mental health treatment, but are never able to receive care (not even an initial assessment session), a rate substantially higher than in the base model ($n = 24$). Furthermore, cumulative mental health enrollment and recovery rates are reduced by approximately 5 % (Fig. 3c).

Service Improvement Models

Service Capacity Approach

To evaluate the impact of increasing system capacity to deliver services in the context of screening, a series of scenarios were modeled in which the universal screening model (above) was expanded to include one or two additional full-time mental health service providers. Figure 4 displays these models and the original single-provider model for comparison. Figure 4a provides a breakdown of provider activities/workload across each scenario. Relative to the one-provider model, the two-provider model simply increases the number of students who are able to receive services. Nevertheless, it appears insufficient to address the school's service needs, as 100 % of the available service capacity is in use for the entire duration of the year in this scenario, while unmet service need remains (see below). It is not until a third provider is added that periods of time emerge during which any of the providers are able to operate below 100 % capacity (represented by the indentation at the top of the rightmost area graph, beginning at week 8). This occurs only after all false positive post-

screening assessment sessions have been cleared from the system.

Furthermore, as can be seen in Fig. 4b, although two providers may be sufficient to maintain a relatively constant number of students in the initial assessment session queue (following the initial screening-related influx), our model suggests that there will still be 21 students waiting for an initial assessment session at the conclusion of the school year (11 of whom will ultimately need mental health treatment and an additional 10 who would enter the alternative pathway). According to our models for this hypothetical scenario, three providers are required to completely eliminate this waitlist. In the three-provider scenario, an initial post-screening spike in students needing mental health treatment and waiting for an assessment session is observed, but reduced to zero by week 7. No students in this scenario are left unserved by the end of the year. Cumulative outcomes across the three scenarios demonstrate predictable increases (because 85 % of each provider's time is devoted to this activity) in mental health treatment capacity with 52, 114, and 142 students receiving mental health treatment across the one-, two-, and three-provider scenarios, respectively. Perhaps most importantly, the number of students expected to demonstrate clinically-significant change improves from 18 to 39 and to 51 across the three scenarios.

EBP Implementation Approach

The second compensatory approach evaluated involved the introduction of an evidence-based intervention for adolescent depression, IPT-A, to be delivered by existing mental health service providers. The IPT-A scenario is presented in Fig. 5 alongside the standard universal screening model, which involved the delivery of usual care services. In our
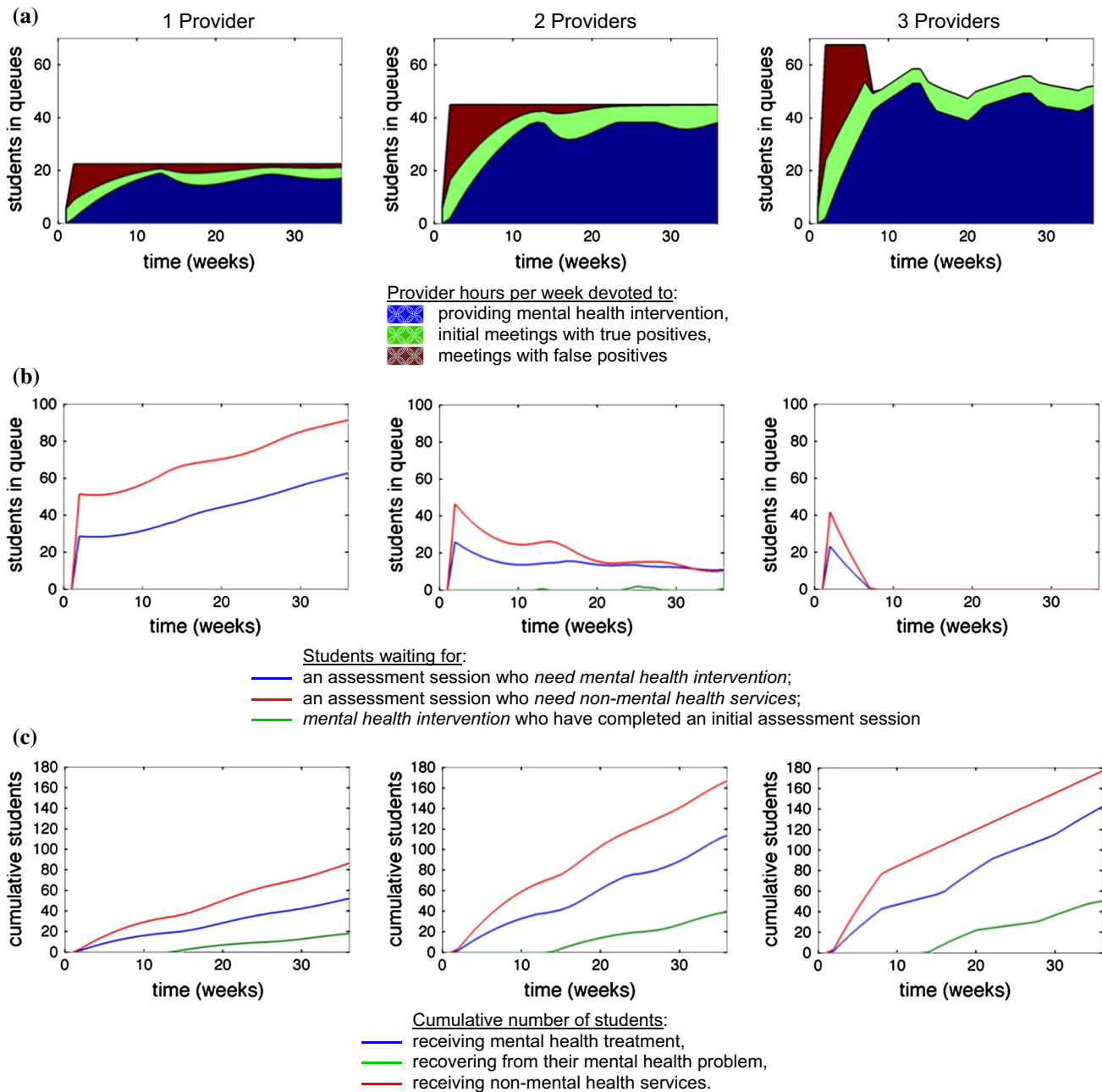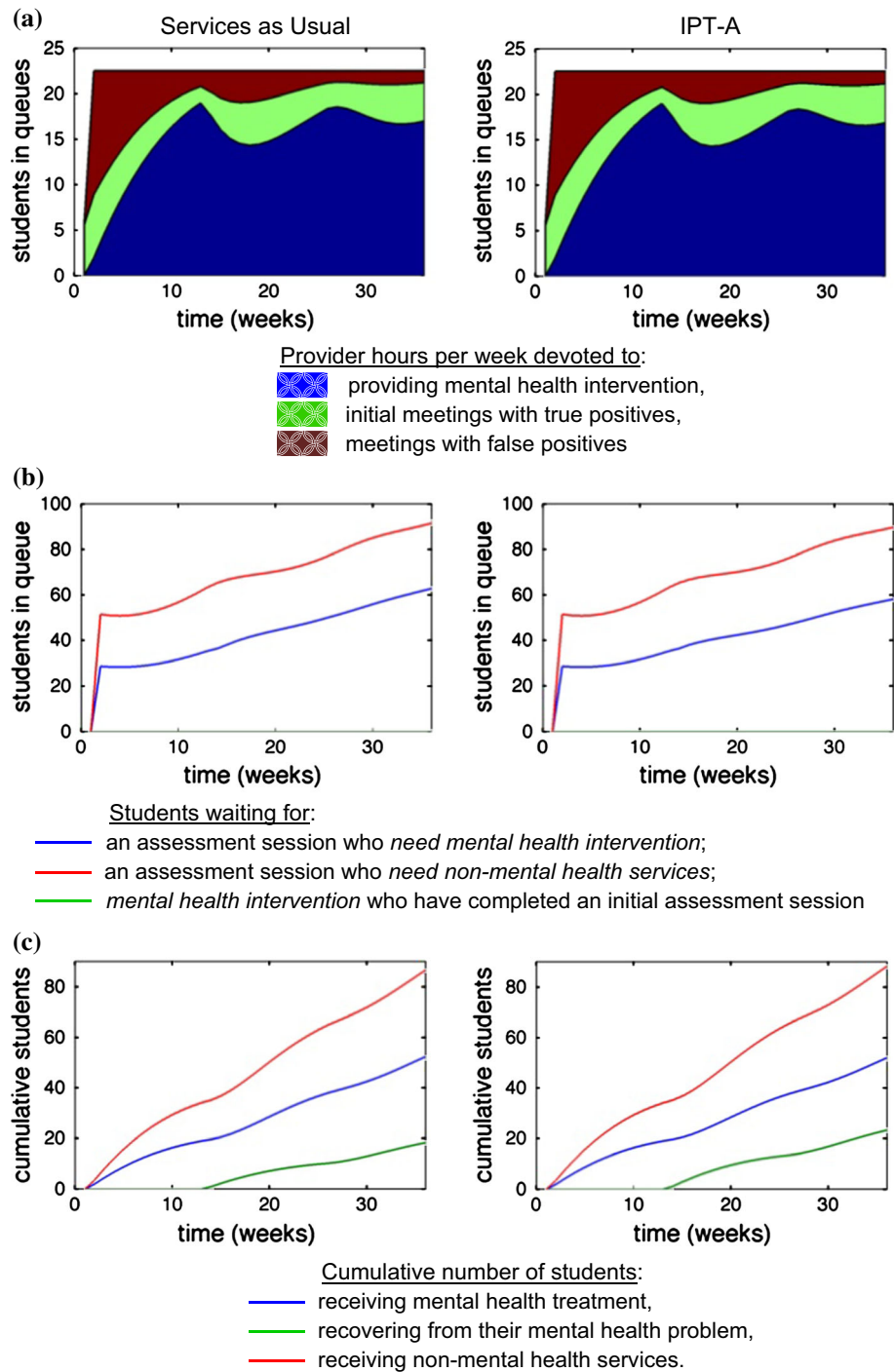
Fig. 4 Service capacity approach with full-time service providers

IPT-A models, the subset of screened students whose initial assessment session confirmed that they were depressed—received the IPT-A intervention.

Comparisons of provider activities (Fig. 5a) reveal virtually no differences between the two scenarios and suggest that providers will have to continue conducting false positive assessment sessions for the entire school year, regardless of whether they are using IPT-A. Figure 5b indicates only slight differences in the number of students requiring mental health treatment and awaiting assessment sessions;

in the IPT-A scenario, the number of students in the queue increases more slowly, with only 58 rather than 63 students waiting at the end of the year. Most notably, although there are few differences between these two models on the number of students receiving mental health treatment, the number of students experiencing a clinically-significant change is somewhat higher in the IPT-A scenario (Fig. 5c). Indeed, it appears that the delivery of IPT-A increases the number of youth who recover from their mental health problem from 18 to 23 over the course of just 1 year.

Fig. 5 Evidence-based practice implementation approach: Interpersonal Therapy for Adolescent Depression (IPT-A)

(a) Services as Usual — IPT-A

Provider hours per week devoted to:
- providing mental health intervention,
- initial meetings with true positives,
- meetings with false positives

(b)

Students waiting for:
- an assessment session who *need mental health intervention*;
- an assessment session who *need non-mental health services*;
- *mental health intervention* who have completed an initial assessment session

(c)

Cumulative number of students:
- receiving mental health treatment,
- recovering from their mental health problem,
- receiving non-mental health services.

Additional Models

Additional models were constructed to probe how to most efficiently address the service needs in the pre-screening and screening scenarios by manipulating components of our universal screening and service improvement models. These models emphasized eliminating student wait queues while ensuring that providers did not spend unnecessary time working below 100 %capacity. Because our previous models suggested that students rarely wait to receive treatment once they have had an initial assessment session, these models focused on eliminating assessment session wait queues. First, models were run to evaluate the impact of including one additional service provider immediately

following screening, dedicated only to conducting post-screening assessment sessions for a 2 or a 4 week period. However, results indicated that this would be insufficient to mitigate the additional need for assessment sessions with students who screened positive on the MFQ and that the number of students in the queue would continue to rise over the course of the academic year.

Subsequent model testing investigated exactly how much provider effort [represented as full time employment (FTE) units] would be necessary to fully eliminate student initial assessment session queues by the conclusion of the academic year across the no-screening, screening, and screening + IPT-A conditions. These models revealed that it would require 2.0, 2.1, and 2.0 FTE, respectively, across the three conditions to eliminate the queue. While the required provider effort is approximately the same, the number of students recovering from their mental health problems increased from 28 to 42 to 51. The first increase from 28 to 42 students is because screening allows for more rapid entry into services and thus more efficient use of available FTE. The second increase from 42 to 51 recovering students is due to the greater effectiveness of IPT-A.

Additional models investigated the role of short-term FTEs devoted to post-screening assessment sessions. Assessment of different combinations of amount of short-term FTE and time periods in the context of the universal screening condition (without IPT-A) indicated that introducing either (1) 1.0 additional FTE devoted only to assessment sessions for 11 weeks or (2) 2.7 additional FTE for 4 weeks represent the most efficient solutions to adequately addressing the short-term service need resulting from universal screening. However, as can be seen in Fig. 6, even in these scenarios, students entering intervention via standard (i.e., non-screening) referral pathways would begin to form a backlog queue soon after the removal of the additional provider FTE (6b). Nevertheless, relative to some other scenarios, these queues are relatively stable, with only 11–12 students who require mental health intervention still waiting for an initial assessment session at the conclusion of the academic year.

Discussion

Nearly 30 years ago, the statistician George Box famously wrote that "all models are wrong, but some are useful" (Box and Draper 1987, p. 424). In keeping with this sentiment, the current paper mined data from the empirical literature which, while imperfect, allowed for the specification of core model parameters and evaluation of our research questions, thus demonstrating the potential utility of a SD approach. This study was designed to contribute to

both the limited literature on applications of SD modeling and the nascent quantitative literature on the impact of universal depression screening on school-based mental health service capacity. For the purposes of this project, a "typical" high school was constructed which contained 1,000 students and one full-time mental health service provider in the base model. Results indicated that, even prior to the introduction of a screening program, estimated unmet service needs of the student population overcame service provider capacity by 10 weeks into the academic year. Results for Research Question (RQ) 1 (what is the impact of universal screening on service need and use?) suggested that, based on our models, screening may generally accelerate the rate of entry into mental health services (including the initial assessment session) for youth, but also had the effect of *slowing* entry into onsite mental health intervention for appropriate students due to the backlog of students requiring an initial assessment session, and ultimately increased the amount of recognized unmet service need. Analyses testing the impact of different compensatory service improvement approaches (RQ 2) revealed that the EBP approach—in which IPT-A was introduced for delivery to depressed youth—had a relatively small impact on service capacity and student outcomes. Findings also suggested that increasing the number of providers in the school to three could be sufficient to address the service need. Nevertheless, additional models indicated that the introduction of short-term providers dedicated to conducting student assessment sessions might most efficiently address the increased service need accompanying the universal screening program. Below, we discuss the implications of these findings for service improvements designed to address and reduce depression in school settings; additional information sources that could facilitate more precise models; limitations of the current study; as well as research and practice applications in service systems.

Service Improvement Implications

Increased Service Capacity

Our results suggested that short-term increases in the number of mental health professionals available to conduct post-screening assessment sessions (i.e., 2.7 additional FTE for 4 weeks) and provide ongoing treatment that is 52 % effective might be the most effective and efficient quality improvement scenario. Although including a single FTE provider for 11 weeks was comparable to using 2.7 FTE for 4 weeks, the relative immediacy of response afforded by more service providers is likely to make this approach more attractive to a range of stakeholders because it would minimize wait times following a positive screen and may
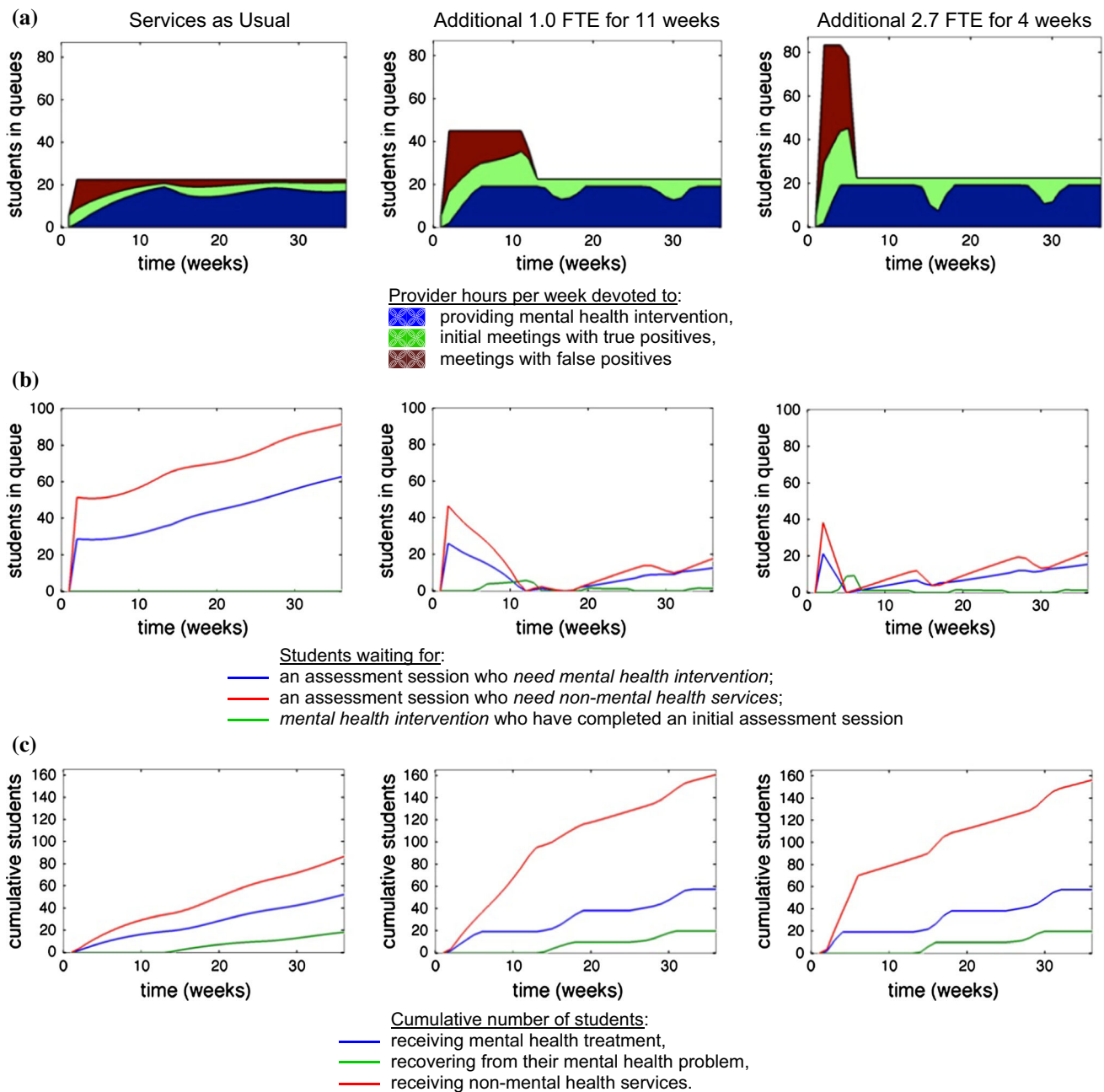
Fig. 6 Impact of time-limited service providers

reduce further disruption of the remainder of the academic year. Even though short-term FTE increases do not completely mitigate the waitlist queue, which begins to grow following their removal, this is likely the most realistic service improvement approach among those tested.

At least one model exists that makes use of child mental health specialists on a short-term basis to conduct 1:1 structured follow-up assessments to confirm or disconfirm screening results and triage students to appropriate academic, social, and mental health supports. In a universal screening program implemented in a large urban school district, performing this time-limited role was attractive to community members with adolescent mental health expertise, providing support for its feasibility (Vander Stoep et al. 2005). Furthermore, it is relatively inexpensive as the costs for the program were found to vary from $8.88 to $13.64 per enrolled student, depending on the prevalence of positive screens in a school (Kuo et al. 2009). This same study found the cost per referral to academic, social, or mental health supports to be $130.22 in schools where

15 % screened positive (slightly higher than the 13.9 % modeled in the current study). Retired school counselors or nurses, part-time community-based therapists, and counseling trainees are cadres that could potentially be tapped for temporary help within this type of model.

EBP Implementation

Interestingly, results suggested that, in our hypothetical models, the introduction of IPT-A had a relatively small impact on the longitudinal and cumulative outcomes. There are two primary reasons why this occurred. First, in our models, IPT-A was only applied in the treatment of depressed youth. Although this was appropriate given that IPT-A is not indicated for other client presentations, it nonetheless limited the impact of the program on the larger service system, in which depressed youth were projected to represent a minority of all service recipients. It is likely that introducing an intervention approach that would be able to address multiple client problems—such as newer modularized, common-elements interventions, which have recently demonstrated superiority to both usual care and traditional evidence-based treatment manuals (e.g., Chorpita et al. 2013; Weisz et al. 2012) and appropriateness for use in schools (e.g., Lyon et al. 2014)—would have yielded a larger system impact. Nevertheless, given that one of the primary aims of the current study was to evaluate approaches specifically designed to compensate for the identification of depressed youth via screening, evaluating a more general service improvement approach exceeded the scope of this project.

Second, despite the fact that IPT-A is among the best-established evidence-based treatments for youth depression (David-Ferdon and Kaslow 2008), we chose a more conservative effectiveness rate from among the range detailed in the IPT-A literature (74 %; Mufson et al. 2004a, b) to replace the identified recovery rate for usual care (52 %). Although IPT-A has been documented to have higher levels of success than those modeled, evidence-based treatments commonly demonstrate attenuated results when implemented outside the context of research studies under less controlled conditions (Burns et al. 2002; Henggeler et al. 1997). Furthermore, multiple studies have documented the difficulties of sustaining interventions with high levels of integrity (Stirman et al. 2012). Given these findings, and data indicating that providing the necessary training and consultation to ensure effective use of EBP is a costly endeavor (e.g., Olmstead et al. 2011), short-term increases in provider FTE may be the more viable option of the two compensatory approaches modeled. This is particularly true given that the IPT-A and short-term FTE scenarios differed only marginally in their impact on student recovery rates (see Figs. 5c, 6c).

Additional Data for Future Model-Building

For the purposes of parsimony and to clearly demonstrate the utility of SD modeling, this paper modeled the impact of a single intervention program on an "average" school. Youth requiring more than one session of mental health services were constrained to just one service delivery pathway with fixed intervention length. However, schools are host to a complex array of academic, health-related, and extracurricular services and programs, characterized by multi-level prevention and intervention approaches, which sometimes span multiple service systems and are delivered by providers with a range of professional backgrounds (e.g., Bohanon and Wu 2011; Brener et al. 2007). There are therefore many additional variables and sources of data that could be incorporated into future SD models. For instance, for simplicity and because external referrals are sometimes unavailable or difficult for providers to successfully facilitate relative to in-school referrals (Lyon et al. 2014), the current models did not incorporate referrals to mental health services external to the school. In addition, although we selected an average rate of students screening positive from the empirical literature, research suggests that the prevalence of positive screens varies across schools (Vander Stoep et al. 2005; Chartier et al. 2008). Each of the variables described above (e.g., treatment length/intensity, multi-level intervention approaches, referrals to outside services, rates of positive screens) represents an example of a model parameter than can be represented as a stock, flow, or feedback loop and systematically adjusted to address additional research questions.

Indeed, there are nearly limitless additional parameters that could be included to more completely reflect the complex school service system context, and the impact of including each parameter can be tested quantitatively. For instance, to evaluate the impact of new depression incidence over the course of the year (our base model only accounted for prevalent cases), we performed additional simulations to include a flow of students who were not initially in need of mental health services, but became depressed over the course of the school year using an annual incidence rate of 7.1/100/year (Lewinsohn et al. 1998). Referring to Fig. 1, an added parameter of incident cases was modeled by adding a flow from the stock of students who do not need treatment to the stocks of students who are depressed and in need of mental health services. This new flow did not significantly impact the number of students who received or recovered from treatment. To promote parsimony, this pathway was not included in our final base model. Further research might specify additional components of the inherently complex school mental health system, such as the effects of turnover and/or timelines for hiring new mental health staff (e.g., exploring the result of spending different amounts of time

posting position advertisements, interviewing applicants, and completing hiring paperwork, orientation and training). Parameters of interest could easily be identified based on the interests of the research team or, more appropriately, the needs of a local school interested in applying a SD approach.

Future models may also examine alternative school-based screening approaches. For instance, Severson et al. (2007) described an assessment approach that relied heavily on teacher nominations to identify at risk youth, which may be a viable option for some youth problem domains. However, our models focused only on student-reported depression symptoms because depression and suicide can be easily overlooked in an adolescent population and are therefore most conducive to self-reported assessment. A 2007 study found that teachers, even with training, nominated fewer than half of students found to be clinically depressed through screening and clinical evaluation (Moor et al. 2007). Another study reported teacher nomination to have a sensitivity of less than 35 % for identifying students with internalizing disorders (Dwyer et al. 2006). Even if teacher-report approaches may be an appropriate alternative, they seem less likely to be effective than models involving student self-report for these reasons.

Finally, despite the wide range of SD modeling possibilities and data sources, much of our ability to model the effects of universal screening is constrained by the information available in the literature. For instance, although there is growing interest in linking school-based mental health programs to educational outcomes (e.g., attendance, academic performance), evidence for this link has been elusive (Pullmann et al. 2013). This is particularly true for school-based screening programs, which have been found to accelerate entry into indicated services (Eklund and Dowdy 2014; Scott et al. 2009), but for which there is little evidence connecting them to changes in meaningful school or academic variables. Beyond its relevance to additional SD models, future research linking universal depression screening to academic success would likely make these programs more appealing and justifiable to stakeholders.

Limitations

The findings of the current study should be interpreted in the context of a few notable limitations. First, although the components and studies selected to populate our models were determined to be the most appropriate for the goals of the current project, they represent only one set of the nearly limitless options and data sources that could be incorporated into such a model. SD modeling is built on the assumption that small changes in a system may have significant, and sometimes unanticipated, consequences for the overall functioning of that system. It is therefore possible that future projects may identify their own set of

relevant studies and come to conclusions that differ from ours in significant and important ways. Second, as indicated earlier, our model was designed to focus primarily on the school context as a service delivery system. As a result, we generated a model that essentially functioned as a "closed system" and did not allow for a flow of students into or out of the school. As a result, external referrals for indicated mental health services were not included (only external referrals for non-mental health services were modeled). In some communities it would be possible for the needs of students whose mental health needs are identified in school-settings to be addressed through referral to community resources. In such communities the results of our study would represent an overestimate of the additional burden of screening on school-based resources.

Practical Applications of System Dynamics Modeling

A number of persistent barriers continue to limit the widespread use of mental health screening programs in middle and high schools (Dever et al. 2012; Fox et al. 2008). Modeling the introduction of a universal screening program and a series of compensatory approaches to address the increase in identified student need enabled us to see how the SD model may serve as a useful step when making decisions about how best to implement new practices. The exercise demonstrated the importance of using such models to directly weigh the pros and cons of introducing a new program and to consider the effects of different variations in program implementation. The current study begins to address one of the most prominent barriers to screening (increased unmet service need) in a hypothetical, average school, but SD modeling could also be used to predict the impact of universal screening and service improvement approaches using data from a specific school. This type of tool would enhance the systematic planning processes schools are already encouraged to use to plan, implement, and evaluate activities (Flaspohler et al. 2012). Recently, Meyers et al. (2012) suggested a set of common "first steps" when any organization considers implementing a new innovation, including assessing current needs and resources and evaluating existing resources. SD modeling clearly complements these first steps. For example, a school could use SD modeling to map existing resources and anticipate problematic gaps in capacity *before* implementing universal screening and/or new interventions. Assumptions could be tested directly with relevant data, leading to more informed problem-solving.

Due to its ability to make use of local data, SD modeling has the potential to produce information that is highly contextually-relevant. Nationwide, there is a growing emphasis on schools' use of local, school-wide data to drive decision-making within a public health framework

(Dowdy et al. 2010; Maras et al. in press). Nevertheless, research suggests school personnel often struggle to effectively use data to plan, implement, evaluate, and sustain their programs (Flaspohler et al. 2008; Roderick 2012). In the context of these efforts, SD modeling could allow school personnel to use local data to "pilot" various hypothetical scenarios based on different configurations of needs and resources to better inform cost-benefit discussions and decision-making. Once developed, SD models could then be used throughout the implementation process to assess service delivery approaches and inform future programming. The adaptability of SD models, and the efficiency with which they can be revised to reflect changing service delivery systems, are among their most appealing attributes.

SD modeling represents a promising approach that could simultaneously enhance and simplify data use in schools. However, additional research is needed to fully realize the practical, on-the-ground promise of SD modeling. While hypothetical, the current study moves us one step closer to that reality. Future research on SD models would contribute to a growing scholarship focused on effective data use in schools. As Roderick (2012) suggests, the time has come to focus less on simply generating or providing data and more on facilitating evidence-based data use.

Conclusion

In sum, the current study demonstrated how SD modeling may have utility in helping to inform new program adoption decisions. Applied to universal depression screening in high schools, our SD models revealed that, consistent with commonly-stated concerns about the drawbacks of screening programs, introduction of the MFQ is likely to further increase the workload of school-based mental health providers. Nevertheless, an approach that makes efficient use of additional, short-term service personnel may be an effective and efficient approach to compensating for this need. Although future research studies should engage in original data collection to substantiate the findings from the current literature-driven models, SD modeling may have a bright future as a viable method for supporting both small- and large-scale implementation decisions across service sectors. Universal depression screening with enhanced school-based service capacity may have a bright future to reduce the burden of depression in adolescent populations.

Appendix

The details of the algorithm that underlies all of the SD models described in this paper are as follows:

1. In the beginning of the school year, the students are partitioned into three stocks (shown as three boxes at the left side of the figure) according to their mental health status.

   a. The *depressed* stock initially contains 139 students, according to data element *a*.

   b. The *non-depressed use MH services* stock is initialized with 142 students, according to data element *b2*.

   c. The *do not need treatment* stock is initialized with the remaining 719 students.

2. If there is no universal screening, then flows to the *true positive* stock will occur only through teacher or self-referrals. These referrals are assumed to begin after the first week and are constant throughout the remaining 35 weeks of the school year.

   a. The flow from stock *b2* to the *true positives* stock is regulated by the valve labeled *referral of others*. The rate of flow is computed by dividing the 142 students in stock *b2* at the beginning of the school year by the time duration of 35 weeks, yielding 4.06 students/week. (Since the SD model is mathematically represented in terms of continuous flows, flow quantities with fractions are permitted. In the results section, when we show the number of students in various queues and treatment states, we will round off to the nearest integer.)

   b. The upper double arrow between stock *a* and the *true positives* stock is for the portion of depressed students likely to use mental health services prior to introduction of the screening program, data element *b1*. As indicated in Table 1, there are 58 students in this category; hence the rate of flow is 58 students/35 weeks = 1.66 students/week.

3. If there is universal screening, then there would be additional flows to the *true positive* and *false positive* stocks. It is assumed that screening occurs during the first week and that the flows to the *true positive* and *false positive* stocks, representing referrals due to screening recommendations, occur during the second week. The rate of universal screening is set by data element *c*.

   a. The lower double arrow between stock *a* and the *true positives* stock is for depressed students who participate in universal screening, and are subsequently referred for mental health services according to the true positive rate. The flow rate is given by the number of students who are depressed and do not fall in category *b1* (139 students in stock *a* minus 58 students characterized by *b1*) times the universal screening rate (data element *c*) times the true positive rate (data element *d1*) divided by the time duration of the flow, or (139 - 58) * 0.825 * 0.71 students/1 week = 47.4 students/week.

   b. The double arrow at the bottom of the figure between the *do not need treatment* and *false positives* stocks is for students who are not in need of mental health services and who participate in universal screening, and are subsequently referred to mental health services according to the false positive rate. The flow rate is given by the number of students who do not need treatment times the universal screening rate (data element *c*) times the false positive rate (data element *d2*) divided by the time duration of the flow, or 719 * 0.825 * 0.29 students/1 week = 172.0 students/week.

4. Students in the *true positive* and *false positive* stocks enter the *MH assessment* and *initial assessment* stocks when they are interviewed by the mental health service providers.

   a. The combined flow into these assessment stocks is given by the number of available interview slots that are available each week. The number of such slots is recomputed weekly, as explained in (8) below.

   b. The proportion of this flow that enter the *MH assessment* or *initial assessment* stock is set to be equal to the proportion of students in the *true positive* or *false positive* stocks.

5. After the students are assessed, they proceed to subsequent stocks in the system, as detailed in the following:

   a. All false positive students return to the *do not need treatment* stock.

   b. All students originating from stock *b2* would flow from the *MH assessment* stock to the *non-MH treatment* stock.

   c. The remaining students, originating from stock *a*, would flow to the *Level 1 treatment queue* or *non-MH treatment* stock. The proportion of students entering the two stocks is given by data elements *f1* and *f2*.

6. The students in the *Level 1 treatment queue* stock will flow to the *Level 1 treatment* stock when a slot becomes available, as determined by (8).

7. Students in the *Level 1 treatment* stock remain in this stock for 12 weeks (data element *g*). After 12 weeks, the students, having completed the Level 1 treatment, flow to one of the following stocks:

   a. The recovery proportion (data element *h1*) flows to the *do not need treatment* stock.

   b. The remission proportion (data element *h2*) flows back to the *true positives* stock.

   c. If evidence-based treatment is used, then the proportion in (7a) is given by data element *i,* and the proportion in (7b) is given by its complement.

8. The number of slots per week that are available for interviews and Level 1 treatment are determined as follows:

   a. The total number of slots per week is given by the number of mental health service providers multiplied by the staff capacity (data element *e*).

   b. Each service provider would allocate 15 % of all slots for interviews and the remaining 85 % for Level 1 treatment.

   c. Additional service personnel would allocate all of their slots for interviews.

   d. If the number of students filling interview slots is less than the total number of slots allocated for interviews, then students in the Level 1 queue can fill the unoccupied interview slots.

   e. If the number of students in the Level 1 queue is less than the number of Level 1 treatment slots, then any students waiting for interviews can fill the unoccupied Level 1 treatment slots.

The above nine-part description of the model constitutes the algorithm that completely governs the flow of students in the system shown in Fig. 1. While this algorithm could be implemented using Vensim equation models, we decided to use MATLAB (MathWorks, Natick, MA) because of its significantly faster computational speed.

References

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental*

Health and Mental Health Services Research, 38(1), 4–23. doi:10.1007/s10488-010-0327-7.

Adelman, H. S., & Taylor, L. (2007). Systemic change for school improvement. *Journal of Educational and Psychological Consultation, 17*(1), 55–77. doi:10.1207/s1532768Xjepc1701_3.

American School Counselor Association (2011). Student-to-Counselor Ratio by State 2010–2011. Retrieved from http://www.schoolcounselor.org/asca/media/asca/home/ratios10-11.pdf. Accessed 14 Jan 2015.

Angold, A., Costello, E. J., Pickles, A., Winder, F., & Silver, D. (1987). The development of a questionnaire for use in epidemiological studies of depression in children and adolescents. Unpublished manuscript.

Banh, M. K., Crane, P. K., Rhew, I., Gudmundsen, G., Stoep, A. V., Lyon, A., & McCauley, E. (2012). Measurement equivalence across racial/ethnic groups of the mood and feelings questionnaire for childhood depression. *Journal of Abnormal Child Psychology, 40*(3), 353–367.

Becker, E. M., & Jensen-Doss, A. (2013). Computer-assisted therapies: Examination of therapist-level barriers to their use. *Behavior Therapy, 44*(4), 614–624.

Best, K. M., Hauser, S. T., Gralinski-Bakker, J. H., Allen, J. P., & Crowell, J. (2004). Adolescent psychiatric hospitalization and mortality, distress levels, and educational attainment: Follow up after 11 and 20 years. *Archives of Pediatrics and Adolescent Medicine, 158*(8), 749–752. doi:10.1001/archpedi.158.8.749.

Bohanon, H., & Wu, M. J. (2011). Can prevention programs work together? An example of school-based mental health with prevention initiatives. *Advances in School Mental Health Promotion, 4*(4), 35–46.

Box, G. E., & Draper, N. R. (1987). *Empirical model-building and response surfaces*. New York: Wiley.

Bradshaw, C. P., Buckley, J., & Ialongo, N. (2008). School-based service utilization among urban children with early-onset educational and mental health problems: The squeaky wheel phenomenon. *School Psychology Quarterly, 23*(2), 169–186. doi:10.1037/1045-3830.23.2.169.

Brener, N. D., Weist, M., Adelman, H., Taylor, L., & Vernon-Smiley, M. (2007). Mental health and social services: Results from the school health policies and programs study 2006. *Journal of School Health, 77*(8), 486–499.

Burns, T., White, I., Byford, S., Fiander, M., Creed, F., & Fahy, T. (2002). Exposure to case management: Relationships to patient characteristics and outcome: Report from the UK 700 trial. *The British Journal of Psychiatry, 181*(3), 236–241.

Chartier, M., Vander Stoep, A., McCauley, E., Herting, J. R., Tracy, M., & Lymp, J. (2008). Passive versus active parental permission: Implications for the ability of school-based depression screening to reach youth at risk. *Journal of School Health, 78*(3), 157–164.

Chorpita, B. F., Weisz, J. R., Daleiden, E. L., Schoenwald, S. K., Palinkas, L. A., Miranda, J., et al. (2013). Long-term outcomes for the Child STEPs randomized effectiveness trial: A comparison of modular and standard treatment designs with usual care. *Journal of Consulting and Clinical Psychology, 81*(6), 999.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implement Sci, 4*(1), 50.

David-Ferdon, C., & Kaslow, N. J. (2008). Evidence-based psychosocial treatments for child and adolescent depression. *Journal of Clinical Child & Adolescent Psychology, 37*(1), 62–104.

De Leo, D., & Heller, T. S. (2004). Who are the kids who self-harm? An Australian self-report school survey. *Medical Journal of Australia, 181*(3), 140–144.

Dever, B. V., Raines, T. C., & Barclay, C. M. (2012). Chasing the unicorn: Practical implementation of universal screening for behavioral and emotional risk. *School Psychology Forum: Research in Practice, 6*(4), 108–118.

Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-based screening: A population-based approach to inform and monitor children's mental health needs. *School Mental Health, 2*(4), 166–176.

Dwyer, S. B., Nicholson, J. M., & Battistutta, D. (2006). Parent and teacher identification of children at risk of developing internalizing or externalizing mental health problems: A comparison of screening methods. *Prevention Science, 7*(4), 343–357.

Eklund, K., & Dowdy, E. (2014). Screening for behavioral and emotional risk versus traditional school identification methods. *School Mental Health, 6*(1), 40–49.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., 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, National Implementation Research Network.

Flaspohler, P. D., Duffy, J., Wandersman, A., Stillman, L., & Maras, M. A. (2008). Unpacking prevention capacity: An intersection of research-to-practice models and community-centered models. *American Journal of Community Psychology, 41*(3–4), 182–196.

Flaspohler, P. D., Meehan, C., Maras, M. A., & Keller, K. E. (2012). Ready, willing, and able: Developing a support system to promote implementation of school-based prevention programs. *American Journal of Community Psychology, 50*(3–4), 428–444.

Fletcher, J. M. (2008). Adolescent depression: Diagnosis, treatment, and educational attainment. *Health Economics, 17*, 1215–1235. doi:10.1002/hec.1319.

Forrester, J. W. (2009). *Learning through system dynamics as preparation for the 21st century*. Acton, MA: Creative Learning Exchange.

Foster, S., Rollefson, M., Doksum, T., Noonan, D., Robinson, G., & Teich, J. (2005). School Mental Health Services in the United States, 2002–2003. *Substance Abuse and Mental Health Services Administration*. DHHS Pub. No. SMA-05-4068. Rockville, MD.

Fox, J. K., Halpern, L. F., & Forsyth, J. P. (2008). Mental health checkups for children and adolescents: A means to identify, prevent, and minimize suffering associated with anxiety and mood disorders. *Clinical Psychology: Science and Practice, 15*(3), 182–211.

Green, J. G., McLaughlin, K. A., Alegría, M., Costello, E. J., Gruber, M. J., Hoagwood, K., et al. (2013). School mental health resources and adolescent mental health service use. *Journal of the American Academy of Child & Adolescent Psychiatry, 52*(5), 501–510.

Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology, 65*(5), 821.

Hirsch, G. B., Levine, R., & Miller, R. L. (2007). Using system dynamics modeling to understand the impact of social change initiatives. *American Journal of Community Psychology, 39*(3–4), 239–253. doi:10.1007/s10464-007-9114-3.

Homer, J. B., & Hirsch, G. B. (2006). System dynamics modeling for public health: Background and opportunities. *American Journal of Public Health, 96*(3), 452–458. doi:10.2105/AJPH.2005.062059.

Hovmand, P. S. (2013). *Community Based System Dynamics*. New York: Springer.

Individuals with Disabilities Improvement Education Act of 2004, Pub. L 101-476, 104 Stat. 1142, codified as amended at 20 U.S.C. § 1400.

Kataoka, S. H., Zhang, L., & Wells, K. B. (2002). Unmet need for mental health care among US children: Variation by ethnicity and insurance status. *American Journal of Psychiatry, 159*(9), 1548–1555.

Kuo, E., Stoep, A. V., McCauley, E., & Kernic, M. A. (2009). Cost-effectiveness of a school-based emotional health screening program. *Journal of School Health, 79*(6), 277–285.

Leischow, S. J., & Milstein, B. (2006). Systems thinking and modeling for public health practice. *American Journal of Public Health, 96*(3), 403.

Levy, D. T., Mabry, P. L., Graham, A. L., Orleans, C. T., & Abrams, D. B. (2010). Exploring scenarios to dramatically reduce smoking prevalence: A simulation model of the three-part cessation process. *American Journal of Public Health, 100*(7), 1253–1259.

Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1998). Major depressive disorder in older adolescents: Prevalence, risk factors, and clinical implications. *Clinical Psychology Review, 18*(7), 765–794.

Lyon, A. R., Ludwig, K. A., Romano, E., Koltracht, J., Vander Stoep, A., & McCauley, E. (2014). Using modular psychotherapy in school mental health: Provider perspectives on intervention-setting fit. *Journal of Clinical Child & Adolescent Psychology, 43*, 890–901.

Lyon, A. R., Ludwig, K. A., Vander Stoep, A., Gudmundsen, G., & McCauley, E. (2013). Patterns and predictors of mental health-care utilization in schools and other service sectors among adolescents at risk for depression. *School Mental Health, 5*(3), 155–165.

Mabry, P. L., Olster, D. H., Morgan, G. D., & Abrams, D. B. (2008). Interdisciplinarity and systems science to improve population health: A view from the NIH office of behavioral and social sciences research. *American Journal of Preventive Medicine, 35*(2 Suppl), S211–S224. doi:10.1016/j.amepre.2008.05.018.

Maras, M., Thompson, A. M., Thornburg, K. R., Hawks, J. S., & Lewis, C. (in press). The transprofessional integration of behavior, academic, and social emotional learning. *Journal of Educational & Psychological Counseling*.

McCormick, E., Thompson, K., Vander Stoep, A., & McCauley, E. (2009). The case for school-based depression screening: Evidence from established programs. *Report on Emotional and Behavioral Disorders in Youth, Fall*, pp. 91–99.

Meadows, D. H. (1999). *Leverage points: Places to intervene is a system*. Heartland, VT: The Sustainability Institute.

Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., et al. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Study-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 980–989.

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology, 50*(3–4), 462–480.

Moor, S., Ann, M., Hester, M., Elisabeth, W. J., Robert, E., Robert, W., & Caroline, B. (2007). Improving the recognition of depression in adolescence: Can we teach the teachers? *Journal of Adolescence, 30*(1), 81–95.

Mufson, L., Dorta, K. P., Moreau, D., & Weissman, M. M. (2011). *Interpersonal psychotherapy for depressed adolescents* (2nd ed.). New York: Guilford.

Mufson, L. H., Dorta, K. P., Olfson, M., Weissman, M. M., & Hoagwood, K. (2004a). Effectiveness research: Transporting interpersonal psychotherapy for depressed adolescents (IPT-A) from the lab to school-based health clinics. *Clinical Child and Family Psychology Review, 7*(4), 251–261.

Mufson, L., Dorta, K. P., Wickramaratne, P., Nomura, Y., Olfson, M., & Weissman, M. M. (2004b). A randomized effectiveness trial of interpersonal psychotherapy for depressed adolescents. *Archives of General Psychiatry, 61*(6), 577–584.

Mufson, L., Weissman, M. M., Moreau, D., & Garfinkel, R. (1999). Efficacy of interpersonal psychotherapy for depressed adolescents. *Archives of General Psychiatry, 56*(6), 573–579.

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.

Olmstead, T., Carroll, K. M., Canning-Ball, M., & Martino, S. (2011). Cost and cost-effectiveness of three strategies for training clinicians in motivational interviewing. *Drug and Alcohol Dependence, 116*(1), 195–202.

Pullmann, M. D., Bruns, E. J., Daly, B. P., & Sander, M. A. (2013). Improving the evaluation and impact of mental health and other supportive school-based programmes on students' academic outcomes. *Advances in School Mental Health Promotion, 6*(4), 226–230.

Roderick, M. (2012). Drowning in data but thirsty for analysis. *Teachers College Record, 114*(11), 1–9.

Romer, D., & McIntosh, M. (2005). The role of primary care physicians in detection and treatment of adolescent mental health problems. In D. L. Evans & E. B. Foa (Eds.), *Treating and preventing adolescent mental health disorders: What we know and what we don't know* (p. 579). New York: Oxford.

Rosselló, J., & Bernal, G. (1999). The efficacy of cognitive-behavioral and interpersonal treatments for depression in Puerto Rican adolescents. *Journal of Consulting and Clinical Psychology, 67*(5), 734.

Scott, M. A., Wilcox, H. C., Schonfeld, I. S., Davies, M., Hicks, R. C., Turner, J. B., & Shaffer, D. (2009). School-based screening to identify at-risk students not already known to school professionals: the Columbia suicide screen. *American Journal of Public Health, 99*(2), 334–339.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology, 45*, 193–223. doi:10.1016/j.jsp.2006.11.003.

Shaffer, D., Gould, M. S., Fisher, P., Trautman, P., Moreau, D., Kleinman, M., & Flory, M. (1996). Psychiatric diagnosis in child and adolescent suicide. *Archives of General Psychiatry, 53*(4), 339–348.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Boston, MA: McGraw-Hill.

Sterman, J. D. (2006). Learning from evidence in a complex world. *American Journal of Public Health, 96*(3), 505–514.

Stirman, S. W., Kimberly, J., Cook, N., Calloway, A., Castro, F., & Charns, M. (2012). The sustainability of new programs and innovations: A review of the empirical literature and recommendations for future research. *Implementation Science, 7*(1), 1–19.

Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2008). *The NSDUH Report - Major depressive episode among youths aged 12 to 17 in the United States: 2004 to 2006*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Tobias, M. I., Cavana, R. Y., & Bloomfield, A. (2010). Application of a system dynamics model to inform investment in smoking cessation services in New Zealand. *American Journal of Public Health, 100*(7), 1274–1281.

U.S. Department of Education Office of Special Education and Rehabilitative Services. (2002). A *new era: Revitalizing special education for children and their families*. Washington, DC.

U.S. Department of Health and Human Services. (1999). *Mental health: A report of the surgeon general*. Rockville, MD: U.S.

Department of Health and Human Services, Substance Abuse and Mental Health Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.

U.S. Public Health Service. (2000). Report of the Surgeon General's Conference on Children's Mental Health: A National Action Agenda. Department of Health and Human Services, Washington, DC. Stock No. 017-024-01659-4 ISBN No. 0-16-050637-9.

Vander Stoep, A., McCauley, E., Thompson, K. A., Herting, J. R., Kuo, E. S., Stewart, D. G., et al. (2005). Universal emotional health screening at the middle school transition. *Journal of Emotional and Behavioral Disorders, 13*(4), 213–223.

Ventana Systems, Inc. (2006). Vensim PLE Software, Ventana Systems Inc. Retrieved May 30, 2014, from the website temoa: Open Educational Resources (OER) Portal at http://www.temoa.info/node/3682.

Walker, B., Cheney, D., Stage, S., Blum, C., & Horner, R. H. (2005). Schoolwide screening and positive behavior supports identifying and supporting students at risk for school failure. *Journal of Positive Behavior Interventions, 7*(4), 194–204.

Walker, S. C., Kerns, S. E., Lyon, A. R., Bruns, E. J., & Cosgrove, T. J. (2010). Impact of school-based health center use on academic outcomes. *Journal of Adolescent Health, 46*(3), 251–257.

Warren, K. (2004). Why has feedback systems thinking struggled to influence strategy and policy formulation? Suggestive evidence, explanations and solutions. *Systems Research and Behavioral Science, 21*(4), 331–347.

Weersing, V. R., & Weisz, J. R. (2002). Community clinic treatment of depressed youth: Benchmarking usual care against CBT clinical trials. *Journal of Consulting and Clinical Psychology, 70*(2), 299.

Weist, M. D. (2008). Enhancing implementation quality. *Advances in School Mental Health Promotion, 1*(3), 2–5.

Weist, M. D., Rubin, M., Moore, E., Adelsheim, S., & Wrobel, G. (2007). Mental health screening in schools. *Journal of School Health, 77*(2), 53–58.

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., et al. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archives of General Psychiatry, 69*(3), 274–282.

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006a). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *American Psychologist, 61*(7), 671.

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006b). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin, 132*(1), 132.

Wilcox-Gok, V., DE Marcotte, Farahati, F., & Borkoski, C. (2004). Early onset depression and high school dropout. In D. Marcotte & V. Wilcox-Gok (Eds.), *Research in human capital and development: The economics of gender and mental illness* (Vol. 15, pp. 27–51). Amsterdam: Elsevier Science.

Williams, S. B., O'Connor, E. A., Eder, M., & Whitlock, E. P. (2009). Screening for child and adolescent depression in primary care settings: a systematic evidence review for the US Preventive Services Task Force. *Pediatrics, 123*(4), e716–e735.

Young, J. F., Mufson, L., & Davies, M. (2006). Efficacy of Interpersonal Psychotherapy-Adolescent Skills Training: An indicated preventive intervention for depression. *Journal of Child Psychology and Psychiatry, 47*(12), 1254–1262.