

Developing Education Implementation and Scaling Capacity

Dean Fixsen, Karen Blase, Caryn Ward, Kathleen Ryan Jackson

National Implementation Research Network

<https://www.activeimplementation.org>

dean.fixsen@activeimplementation.org

Abstract

This report documents the culmination of ten years of development of a new approach to improving education systems, organizations, and outcomes. Based on the Active Implementation Frameworks, intensive support is provided to develop implementation and scaling infrastructures in state education systems to initiate and manage change processes, and to provide reliable supports for improved teacher instruction and student learning. Measures of capacity monitor progress in states and inform action planning. The results from 5 states in Study 1 were encouraging and provided the opportunity to engage in improvement activities for capacity development methods and measures. The results from 5 states in Study 2 provide evidence of consistent and relatively rapid progress in capacity development replicated across states. Implementation practice and science are significant additions to efforts to improve education in the United States.

Keywords: implementation, scaling, improvement, capacity, system change, education

Developing Education Implementation and Scaling Capacity

The goal of education is “to promote student achievement and preparation for global competitiveness by fostering educational excellence and ensuring equal access” (U.S. Department of Education, 2009). Developing the capacity to achieve this mission is the responsibility of state, regional, and local education agencies. In education, “capacity” is defined as, “the perceived abilities, skills, and expertise of school leaders, teachers, faculties, and staffs [to] accomplish something specific, such as leading a school-improvement effort and supporting teacher’s effective use of a practice so students have equitable access to practices that do improve their educational outcomes. The term may also encompass the quality of adaptation—the ability of a school or educator to grow, progress, or improve.”

(<http://edglossary.org/capacity>). The purpose of this article is to describe a new approach to developing implementation capacity in state education systems, introduce a measure of state capacity, and describe data from replications of a capacity development process across ten states.

Implementation Capacity Development

Implementation capacity development necessarily focuses on processes to align organization and system structures, roles, and functions in a state education system. Functions consist of the knowledge, skills, and abilities of individuals who perform particular roles within system structures. The new approach to capacity development is based on implementation science (Ben Charif et al., 2017; Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Saetren, 2014).

“Implementation science” is defined as, “The study of factors that influence the full and effective use of innovations in practice” (Fixsen, Blase, Metz, & Van Dyke, 2015). Implementation is the “to” in science to service, research to practice, policy to practice, innovation to outcome, and so on. “Implementation capacity” is defined as “the availability of implementation teams with the knowledge, skills, and abilities to develop competencies and affect positive change with practitioners, organizations, and systems” (Fixsen, Blase, & Fixsen, 2017, p. 488).

Implementation: the Missing Link

Effective innovations and enabling policy contexts have been the focus of education reform efforts for several decades without producing significant, sustainable, and scalable benefits to students (Elmore, 2002; Manna, 2008; National Center for Education Statistics, 2013; National Commission on Excellence in Education, 1983; Tyack & Cuban, 1995). The lack of attention to effective implementation may account for the lack of educationally significant outcomes. For example, Vernez, Karam, Mariano, and DeMartini (2006, p. iii) stated “the effect of comprehensive school reform models on student achievement remains debatable. Research results have been mixed. Most studies show only a modest effect—or sometimes no effect—on student achievement. One important reason for this mixed record is that most prior studies have not accounted for the extent to which schools have actually implemented their adopted models.” In their study of 8,000 schools supported by over \$2 billion to use a research-based

comprehensive school reform (CSR), Vernez et al. (2006) found minimal support provided for teachers' professional development, little support for staff time from principals, little use of the chosen CSR as intended, and no impact on student outcomes. The authors found that "schools tended to engage in the same types of activities regarding curriculum, methods of instruction, student groupings, governance, assessment of students, and parent involvement regardless of whether the school used one of the four models or not" (p xx). Vernez et al. concluded, "If comprehensive school reform has not been implemented, or has been implemented only in part, changes in student achievement cannot be expected—or, if such changes occur, they cannot necessarily be attributed to the reform" (p iii).

Poor outcomes that likely are due to a lack of implementation supports continue to be replicated in the federal government's investment of over \$3 billion in School Improvement Grants (SIG). Recent analyses revealed no significant impact on math or reading achievement outcomes for students in any grades, high school graduation rates, or college enrollment for schools using a SIG-funded model. The SIG program had no discernable impact on SIG-funded practices in schools (Dragoset et al., 2017). Thus, the outcomes of SIG models are unknown because the SIG models were not used in practice.

In contrast, the Positive Behavior Intervention and Supports (PBIS) whole school intervention includes many aspects of implementation best practices and has been scaled to over 25,000 schools (Horner, Sugai, & Fixsen, 2017). Fidelity of PBIS use in schools is regularly assessed and fidelity scores are significantly related to student outcomes (Horner et al., 2004; McIntosh, Mercer, Nese, & Ghemraoui, 2016). Thus, the problem is not that the innovations are ineffective in the Vernez et al. (2006) and Dragoset et al. (2017) examples; the problem is that potentially effective innovations were not used as intended in practice and, therefore, students did not benefit from innovations they did not experience.

Capacity Development and Systemic Change

The need to strengthen the capacity for change in education has been noted for several decades (e.g. Adelman & Taylor, 2003; Darling-Hammond & McLaughlin, 1995; Freeman et al., 2010; Hall & Hord, 1987; Manna, 2008; National Commission on Excellence in Education, 1983). In spite of multiple national and state reforms and the evidence in support of practices that improve learner achievement (Hattie, 2009), the National Assessment of Educational Progress (NAEP) literacy scores for students at age 9 have hovered around a mean of 215 on a 500-point scale since the 1960s. In 2006, the importance of implementation science as an element of capacity was recognized by the U.S. Department of Education Office of Special Education Programs (OSEP) and led to the development of what is now the State Implementation and Scaling up of Evidence-Based Programs (SISEP) center.

A goal of the SISEP center is to make use of the best available evidence in implementation practice, science, and policy to develop capacity to support meaningful change in state education

systems. OSEP and SISEP recognized from the beginning that any effort to improve education outcomes for all students essentially is a system change effort (Barber & Fullan, 2005; Fullan, 2001). Systems have multiple layers of moving parts that are connected in many formal and informal ways. While planning is important, there really is no *a priori* way to “analyze a path” to system change (Beyer & Trice, 1982). To paraphrase Beyer and Trice (1982), merely thinking about system change will not take us very far. However, there is a way to “behave a path” to transformative system change (Ulrich, 2002). Once the activities involved in developing implementation capacity begin, the activities disturb existing relationships in the system (Marzano, Waters, & McNulty, 2005; Morgan & Ramirez, 1983; Nielsen, 2005). The results of those disturbances reveal apparent and previously unknown connections and lack of connections among system components (Conklin, 2005). Previously unknown proponents and detractors suddenly appear and exert influence (Rittel & Webber, 1973). As soon as the reactions are known, actual facilitators can be strengthened and relevant impediments can be resolved (Marzano et al., 2005; Ulrich, 2002). In this way, behaving a path to system change, also known as “enactment in practice” (Svensson, Tomson, & Rindzeviciute, 2017), improves focus and improves efficiency and effectiveness of the change process.

OSEP SISEP Implementation-Informed Plan

In 2007 OSEP and SISEP conceptualized the development of an implementation infrastructure (structure, roles, and functions) to support the full and effective use of evidence-based practices in schools. The “implementation-change” structure to be developed in each state is shown in Figure 1. The structure is based on the best evidence from implementation, organization, and complexity science. From these points of view, state education systems are “nested systems” with sub-systems (regions, districts, schools) that have their own logic, conditions and specific functions in relation to the others (Resnick, 2010). The “loosely coupled” sub-systems (Hord,

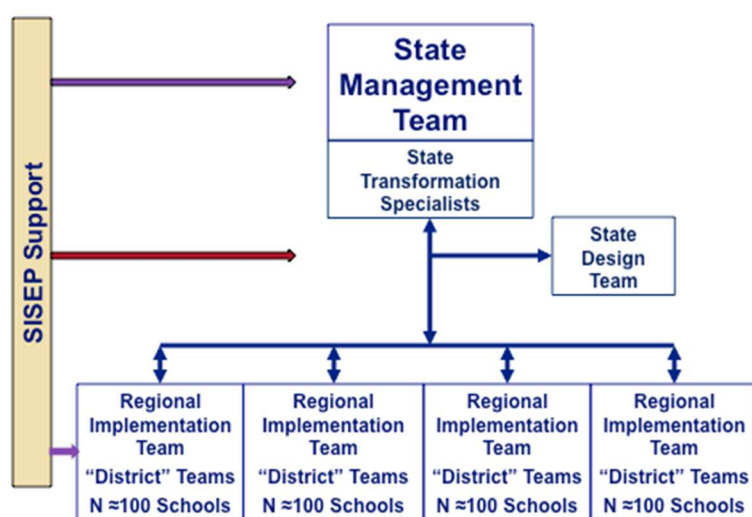


Figure 1. The OSEP SISEP plan for scaling effective innovations in state education systems.

1992; Zucker, 1987) relate to one another in an “environment of turbulence, flux, fragmentation, disequilibrium and uncertainty” (Brachthausen, 2011, p. 222) that makes planned change difficult and unsustainable. To account for this complexity, implementation-informed system change efforts can develop “scaled redundancy” (fractals; Kluger, 2008) in the form of linked implementation teams, and build in “recursive feedback loops” (Gilpin & Murphy, 2008) to help assure prompt top down

support for bottom up change (Darling-Hammond & McLaughlin, 1995; Sabatier, 1986). In a state education system, district and school subsystems are more tightly coupled than state-region-district subsystems. Some districts are more tightly coupled than others and vary in size but even large districts are more tightly coupled with their schools than a state is with its regions and districts. The design for implementation and scaling capacity development must accommodate the variations in linkages and size of subsystem within and across state education systems.

As shown in Figure 1, the infrastructure (a cascading system of supports) to assure effective implementation supports for teachers and staff in schools consists of:

- The State Management Team (SMT): The Chief State School Officer and his or her cabinet (an existing group) to promote and provide leadership and political visibility for capacity development, alignment, and integration of State Education Agency (SEA) priorities in support of districts and schools.
- State Transformation Specialists (STSs): Two people selected to lead the capacity development, system change, and alignment work in the SEA, regions, districts, schools, and greater community.
- A State Design Team: Division or department leaders across the SEA who develop or design the agency's common processes and procedures for implementation across priorities or initiatives as part of the capacity development and system change process
- Regional Implementation Teams (RITs): Members selected by a Regional Education Agency to learn implementation best practices in order to develop implementation teams in several districts and schools (multiple districts that total about 100 schools in a region)
- District Implementation Teams (DITs): District executive leadership and staff who are selected to learn implementation best practices to support schools and teachers directly to change school management routines and improve use of effective innovations for instruction to improve student outcomes.
- Building Implementation Teams (BITs): School staff members who are selected to provide direct support to teachers and help change school management routines to improve support for the use of effective innovations and instruction to improve equitable access to practices and student outcomes.

At the district and school level, effective and sustainable implementation supports focus on teacher and staff use of effective education methods (e.g. evidence-based practices; innovations) that produce noticeable student gains. Depending on the size, complexity, and resources of districts and its schools, the district and school implementation teams may be highly integrated (multiple roles and functions are shared by team members at district and school levels).

As shown in Figure 1, the linked implementation teams at state, region, district, and school levels are intended to provide the foundation for scaling innovations across a state education system.

The linked implementation teams each learn to use the Active Implementation Frameworks (Blanchard et al., 2017; Fixsen, Blase, Metz, et al., 2015; Fixsen et al., 2005). The evidence-based Active Implementation Frameworks include usable innovations, implementation drivers, implementation stages, improvement cycles, implementation teams, and systemic change. The Active Implementation Frameworks provide a common language, common methods, and common measures that promote communication among teams. The intended outcome is to align and integrate current structures, roles, and functions so that system resources can be leveraged to maximize supports for teachers' use of effective education practices when interacting with students. Implementation Teams make use of evidence from research and practice to conduct implementation-informed work as part of the roles and functions in region, district, and school structures so that teachers have adequate support for effective instruction when interacting with students in the classroom. Through this process, the different agencies in the system change their "ways of work" to increase their organizational capacity. In this design, the Implementation Team roles and functions are fulfilled by the capable people already employed in state departments of education and in the regional and local education agencies together with stakeholders, families, and community partners. It is the role of SISEP and the STSs to help organize Team members at each level and support them in their capacity to use the Active Implementation Frameworks with expertise and fluency to address the unique needs and circumstances within and across organizational structures.

Transformation Zone

Initially, capacity is developed in a "transformation zone" (Fixsen, Blase, & Van Dyke, 2012), also known as a zone of proximal development (Moll, 1990). In the transformation zone changes are initiated, elements of the existing system are disturbed, and the process of behaving a path to system change is begun. For SISEP, a transformation zone is a vertical slice of the entire education system, from the classroom to the capitol. The zone includes the state leaders and staff, 2 or 3 regions, 3 districts in each region, 3 schools in each district, and all the teachers, staff, and students in each school. A transformation zone is large enough to encompass nearly all the elements of the entire system and small enough not to be overwhelmed by the issues encountered as changes are initiated and managed. The purpose of the transformation zone is to provide opportunities for purposeful training and coaching to establish the first examples of Implementation Teams at each level and have them begin to function independently and together. A related purpose is to disturb the existing system (Marzano et al., 2005) and evoke reactions that provide opportunities for the state management team to make changes in the system to accommodate and support implementation team development and functioning. The practice-policy communication cycle (Fixsen, Blase, Metz, & Van Dyke, 2013) provides bottom up information to improve top down support for implementation and scaling capacity development (recursive feedback). Over time, additional cohorts of regions, districts, and schools are added to the systems change effort as the STSs and colleagues develop an implementation team in each region of the state and each RIT develops implementation teams in each district in their region

and so on to reach every school in a state. In the SISEP OSEP plan, scaling improved student learning is based on increasing the number and quality of linked Implementation Teams.

Intensive Implementation Informed Support

As noted on the left side of Figure 1, it is the role of SISEP and the State Transformation Specialists (STSs) to help organize Team members at each level and help them develop the required Active Implementation expertise. This is not a typical “train the trainer” approach where the original trainers bow out of the process after the first hand off. In Figure 1 and in practice, SISEP staff provide the original training and coaching and assessments at each level and persist with implementation-informed coaching until the Active Implementation frameworks are in use as intended by the implementation team at each level and at all levels simultaneously. From the first contacts with a state, the SISEP staff work with leaders and staff at every level to assure readiness for capacity development and engagement in the processes of change and managing change. In this way, SISEP and colleagues in state systems work together to assure a good fit between implementation team development and local readiness and resources.

Research Aim

The aim of this study is to evaluate the viability of the OSEP SISEP implementation informed plan. According to the plan, implementation and scaling capacity is established at the state level so that additional capacity can be established in regions, districts, and schools to support improved classroom instruction. In this article the focus is on evaluating the impact on education system functioning at the state level. The research was conducted in two studies.

The data presented are the results of repeated measures of capacity development at the state level. Future articles related to Study 2 states will focus on the development and assessment of capacity at the regional (St. Martin, Ward, Harms, Russell, & Fixsen, 2015), district (Ward et al., 2015), school (Fixsen, Blase, Naoom, et al., 2015), and teacher (Fixsen, Ward, Ryan Jackson, & Chaparro, 2015) levels and on the relationship between implementation and scaling capacity at these levels and student outcomes.

Study 1: (2007-2012)

Methods

Participants

In Study 1 a total of 5 state education agencies and a representative sample of their regional and local education agencies actively participated in capacity development. Study 1 states were identified through a mutual selection process that commenced in December 2007 and concluded in May 2008. Mutual selection means that the state leaders and stakeholders make an informed decision to participate and SISEP makes an informed decision regarding a state’s ability to participate. Because a specific focus on implementation was new to educators, the selection

process began with widely advertised invitations to participate in two group phone calls. Representatives from 35 states participated in the calls to hear about implementation and scaling, requirements for participation, and the role of SISEP in helping states develop capacity. Those who were interested in pursuing the work in their state were invited to contact SISEP. The 16 states that responded were sent an email that outlined the state selection criteria and were asked to describe activities in the state related to each criterion. The state selection criteria were:

- The State has documented attempts to implement an evidence-based curriculum or instruction initiative or behavior support program
- The State has demonstrated leadership commitment to system change at the State level
- Within 12 months, the State is willing to allocate current funds and staff positions for two State Transformation Specialists
- By the end of Year 2, the State is willing to allocate current funds and staff positions for Regional Implementation Team members.
- The State is willing to establish a data system that includes assessment of adult (teacher, staff, Implementation Team) behavior as well as student outcomes.
- The State is willing to participate in and contribute to a national community of practice.

The SISEP team and consultants rated the information provided in the 10-page applications from 16 states. Eight states were scored as partially or fully meeting criteria on each selection

Table 1. Study 1: Ratings of information in state applications.

Criterion	HIGHER (n=8)	LOWER (n=8)
EBP	1.77	1.15
Leadership	1.65	0.96
Resources*	1.60	1.11
Data System	1.58	1.19
Comm of Practice	1.38	0.88

Scores: 2 = Fully meets criteria, 1 = Partially meets criteria, 0 = Not currently meeting criteria

* Resources includes ratings of information regarding STS and RIT positions (criteria 3 & 4)

category. The remaining 8 states were scored as not currently meeting criteria on one or more of the selection criteria. The SISEP ratings for the higher and lower rated states are shown in Table 1.

The higher rated states were contacted and site visits were arranged to meet with the state education leaders for up to two hours in the morning and state education stakeholders (e.g. representatives of superintendent and principal associations, parent advocacy groups, teacher unions or associations) for up to two hours in the afternoon. Two members of the SISEP team visited each of the 8 states. During the on-site visits the leaders and stakeholders in 6 states provided further evidence of their current work and enthusiastic leadership and stakeholder support for engaging in implementation and scaling activities. These states were invited to participate as active scaling states. Subsequent to the invitation, lawsuits over which bodies controlled education in one state led to a delay in that state's participation. In another state leadership and political changes led to withdrawal. Four states began in September 2008 and one (delayed) state began in 2011.

Information regarding the Study 1 (n =5) states is provided in Table 2. The states were located in the western, northwestern, midwestern, and southeastern parts of the United States. The states varied in size (120 – 850 districts) and complexity and provide a good test of the OSEP SISEP capacity development plan.

Table 2. State Participants in Study 1.

Study 2	Location	Number Districts
State #1	West	200
State #2	Midwest	850
State #3	Northwest	590
State #4	Northwest	340
State #5	Southeast	120

Intervention

To carry out the OSEP SISEP plan, SISEP provides intensive implementation-informed support for each state (Chinman et al., 2008; Fixsen, Blase, Duda, Naoom, & Van Dyke, 2010; Fixsen, Blase, Horner, & Sugai, 2009). SISEP visits each state each month for 2 or 3 days and engages in telephone conferences and email exchanges between visits. During each visit, SISEP provides training and coaching on the use of the Active Implementation frameworks to develop

implementation teams and to use the experience to change the education system to support implementation capacity and effective education practices more fully. SISEP works directly with the State Transformation Specialists (STSs) and a State Design Team and then coaches the STSs and Design Team members as they work with regional groups to form Regional Implementation Teams. SISEP schedules meetings with the State Management Team (SMT; State Chief School Officer or his/her deputy and the cabinet) during each visit to provide information about implementation and scaling and the progress being made in the state.

Measures

Measures to assess the presence and strength of capacity are required so that capacity can be developed on purpose and changes can be monitored. Measures of capacity have been developed and used in fields outside education (Acosta et al., 2013; Chaple & Sacks, 2016; Flatten, Engelen, Zahra, & Brettel, 2011; Goodman, McLeroy, Steckler, & Hoyle, 1993; Gotham, Brown, Comaty, McGovern, & Claus, 2013; Jiménez-Barrionuevo, García-Morales, & Molina, 2011; Lee & Cameron, 2009; McGovern, Matzkin, & Giard, 2007) and within education (Elmore, Forman, Stosich, & Bocala, 2014). These measures have provided a good foundation for assessing general capacity.

To focus on key implementation factors, the State Capacity Assessment (SCA; Fixsen, Blase, Duda, & Horner, 2009) was developed during the first year of work in Study 1 states.

Table 3. State Capacity Assessment (SCA) and Subscales

State Capacity Assessment Subscales	Concept Areas for Items (46 items)
State Management Team (SMT)	
12 items that address key roles and functions	
State Design Team	
12 items that address key roles and functions	Leadership team, coordination, authority, policy, funding, visibility and political support, evidence-based practice or innovation, system integration, operations, action planning, evaluation, implementation
Regional Implementation Team	
12 items that address key roles and functions	
Evaluation	
10 items that address data collection, reporting, and use	

In Study 1 the SCA consisted of 46 questions that were administered by SISEP in a facilitated meeting (see Table 3). Respondents included the state transformation team made up of the STSs and the state design team members, all of whom were directly involved in the work of developing implementation capacity in the state. An example of an item that assesses roles is, “Leadership Team: The SMT consists of the State superintendent/ commissioner and other elected or appointed leaders who make state-level policy and fiscal decisions.” An item to assess function is, “Authority: The SMT has officially endorsed the SDT [State Design Team] and established the extent of the authority of the SDT to make organizational and fiscal decisions with regard to the infrastructure for implementation of evidence-based practices and other innovations statewide.” Each item on the SCA is scored on a three-point scale where 2 = fully in place, 1 = partially in place, and 0 = not yet in place. SISEP asked respondents to meet to discuss and score each item as a group. The SCA was completed twice a year in each state, typically in the middle of the fall term of an academic year and again in the spring. The SCA score is presented as a percentage of the total possible score of 92 (46 items with a maximum score of 2).

Results

Figure 2 shows the SCA scores for up to 5 years for each state. The trends indicate some progress was made in capacity development in the states participating in Study 1 but was inconsistent across states over 5 years. SISEP had set a benchmark at 60% for “acquisition” of implementation knowledge, skills, and abilities and at 80% for “proficiency.” These benchmarks

are noted in Figure 2. Engagement in implementation and scaling capacity development ended by mutual agreement in State #2 after 23 months and in State #3 after 20 months. SCA scores in those 2 states were under 30% for each administration. SCA scores in State #1 remained low for all 5 years while SCA scores in States #4 and #5 improved over 5 years. The SCA scores reached the 60% acquisition benchmark in these 2 states and sustained for 2 consecutive administrations for State #4.

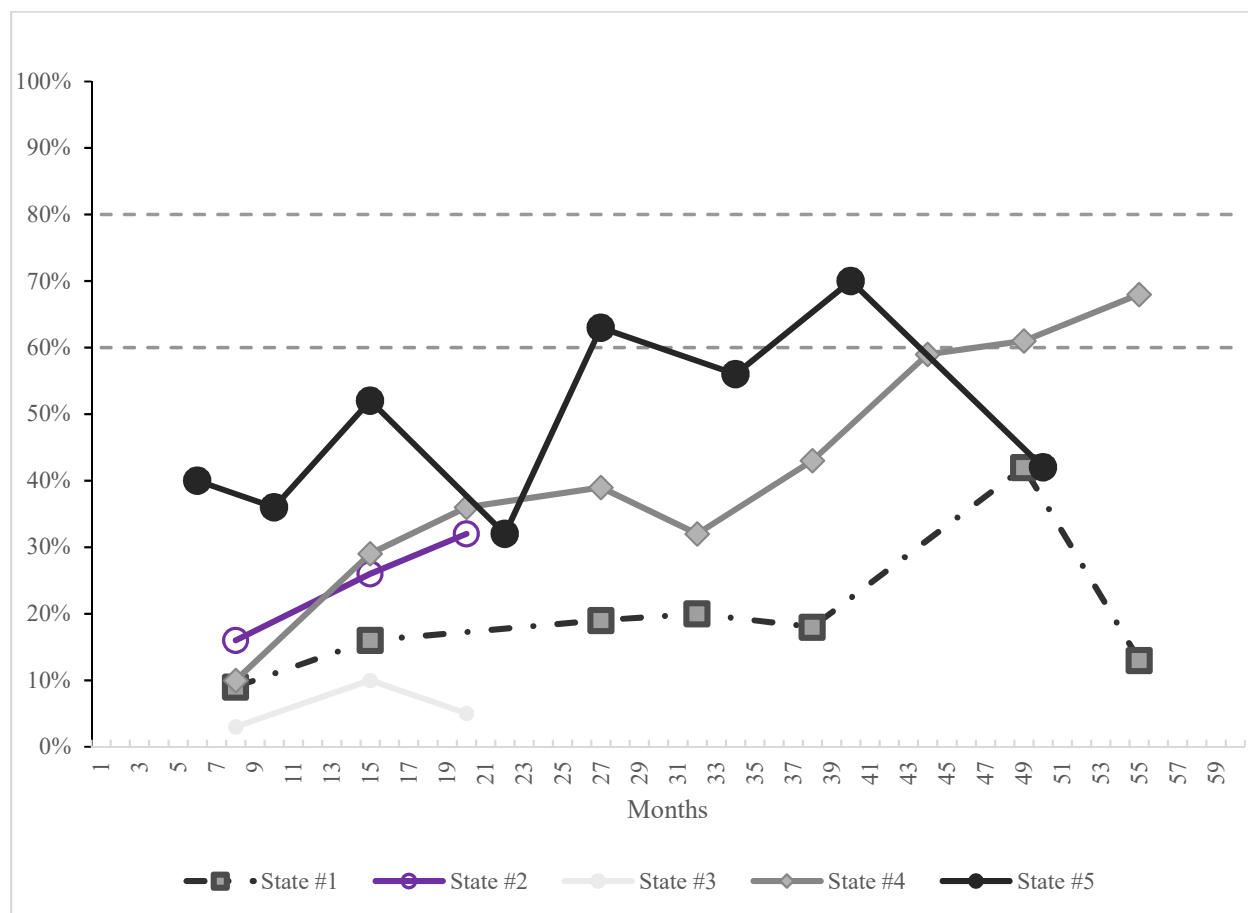


Figure 2. Study 1 State Results on State Capacity Assessment. Month 1 is the month SISEP and state agreed to begin working together to develop state implementation and scaling capacity.

Discussion

The results from Study 1, while mixed, were still encouraging. Staff in state education systems could be engaged in the intensive implementation and scaling capacity development work and 3 of the 5 states were willing to persist in the work for 5 years. For 2 states there were marked improvements noted in repeated SCA assessments. Repeated assessments over 5 years in State #1 were not associated with notable improvements. Thus, from a measurement point of view, reactive measurement effects (Singleton, Straits, Straits, & McAllister, 1988) likely did not

unduly impact SCA scores in state education systems. The learning from Study 1 was applied in Study 2.

Study 2 (2014-2017)

Methods

Participants

In Study 2, 5 states were selected to participate in SISEP following a revised mutual selection process. Two states began in Study 2 in 2014 and one state began in 2015, 2016, and 2017 respectively for a total of five states in Study 2. Study 2 states were mutually selected for participation through the use of revised selection criteria (noted below) based on Study 1 experience, and an exploration process consisting of a series of informational conference calls and site visits to assess goodness of fit, readiness, and leadership commitment for systemic change. Two other states that requested assistance did not meet SISEP selection criteria related to having regional education agencies as part of the education system and were not selected for Study 2. It is noteworthy that State #3 (Study 1) and State #9 (Study 2) are the same state. While it is the same geographic state, State #9 in Study 2 (selected in 2016) has an entirely new leadership group for education compared to State #3 (selected in 2008; ended in 2010).

Study 2	Location	Number Regions	Number Districts
State #6	West	9	300
State #7	Midwest	12	175
State #8	Northeast	12	450
State #9	Northeast	56	590
State #10	Midwest	8	180

Intervention

SISEP maximized the learning from the failures and marginal successes in Study 1 states (Bryk, 2016; Firestein, 2016; Frick, Elder, Hebb, Wang, & Yoon, 2006; Harrison & Grantham, 2018;

Nielsen, 2005; Petroski, 1982). The SISEP OSEP partnership with states since 2007 is an example of usability testing (Akin et al., 2013; Epstein & Klerman, 2013). As shown in Figure 3, a plan is followed with a small group of organizations, the results of doing the plan are intensively studied, and a new and improved plan is initiated with another small group. The resulting “virtuous circle” (Fox & Gershman, 2000) creates a learning organization (Morgan, 1997; Senge, 2006) where improved practice leads to better outcomes and continually improved

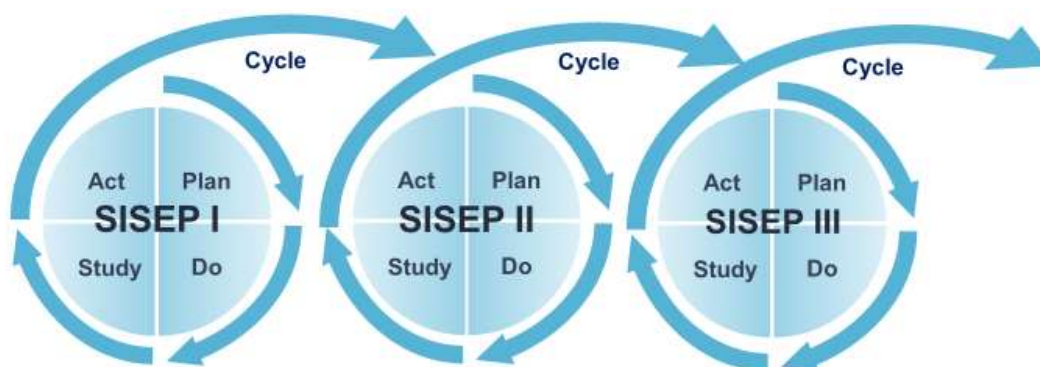


Figure 3. Usability testing employed by SISEP as a learning organization.

plans as a result of each cycle. In this case, the 5 states in Study 1 participated in the SISEP I cycle. That experience was studied and actions were taken to create a new and improved plan for Study 2 (SISEP II) and the analytic process will be repeated with future states in SISEP III. By doing the work of state system change and capacity development on purpose, the SISEP supports for states can be improved on purpose. With usability testing in mind, specific lessons from the Study 1 states included:

- The State Management Team (SMT) needs to be fully engaged and able to meet with SISEP and the STSs at least once each month with one hour on their standing agenda. There was little or no access to the SMT in States #2 and #3 in Study 1 and change efforts were constrained once it became clear that developing capacity required systemic change and SMT engagement in the change process.
 - In exploration and subsequent work in Study 2 states, the CSSO and Cabinet members need to be included in the decision to proceed with implementation scaling capacity development and agree to actively participate in systemic change work.
- The State Transformation Specialists (STSs) need to be fully employed by the state system and reside in the offices of the Chief State School Officer (CSSO). STSs who were employed outside the system (State #1) had no access to formal internal communications (e.g. Outlook calendars, regular meetings) and informal communications (e.g. hallway conversations; pick up meetings), cumbersome access to internal data systems (e.g. data sharing agreements with de-identified data), and no authority to call

meetings or independently approach units within the education system. In States #2 and #3 STSs who were segregated in particular units (e.g. Special Education; school turnaround division) overcame some of these limitations but lacked authority and ready access to other units in the broader system (e.g. literacy, math, behavior, social-emotional learning; data management and reporting).

- In Study 2 states, the requirement to employ and fund at least 2 full time STSs was included in exploration discussions and subsequent action.
- Two SEA sponsors from the Cabinet were also required to provide direct access and support of STSs.
- Regional Education Agencies need to be active partners in the state system and fully engaged in implementation and scaling efforts. States #1, #2, and #3 in Study 1 had existing regional agencies but they were not engaged, in part due to the lack of SMT engagement and support. States #4 and #5 had loosely-configured regional groups of some kind, but the regions had no meaningful role in supporting academic or behavior programs in districts and schools. States #4 and #5 attempted, with moderate success, to develop functional regional agencies and implementation teams in Years 4 and 5 of Study 1.
 - In Study 2 states, the existence of viable regional education agencies and the willingness of the state and region to develop partnerships needs to be verified during the exploration discussion and acted upon expeditiously once work begins in a state.
- In 2007, the original design of SISEP was based on assumptions that a) available evidence-based education practices are in use in every state and most districts, b) a fidelity measure is available to assess the presence and strength of any evidence-based practice in use, and c) data systems in states and districts include measures of education capacity and processes related to producing high levels of student learning. The purpose of SISEP, then, was to strengthen the capacity of state, regional, and district teams to provide supports to schools and teachers based on the latest developments in implementation science. These assumptions were rarely met in Study 1. Yet, the evidence from implementation science is strong that these factors still must be in place in order to produce educationally significant outcomes.
 - In Study 2, the SISEP and STS tasks need to be expanded to include working with state teams to identify and operationalize effective innovations, establishing fidelity measures, and developing measures of education capacity and processes that can be used in state education systems. Knowledge, skills, and abilities to accomplish these tasks need to be built into the development of implementation teams in each state.
- A cascading system of supports, consisting of linked implementation teams, is needed to develop and scale implementation capacity. In 5 years none of the states in Study 1 were able to produce regional or district implementation teams that could meet

“acquisition” standards set by SISEP (a score of 60% on an implementation capacity assessment).

- In Study 2 the teaching and learning in states needs to be parsed into useful lessons that can be put into practice almost immediately to facilitate learning, coaching, and outcomes for capacity development.
- Developing implementation capacity is nearly cost-neutral from a funding standpoint. Every education system has existing staff assigned to improvement activities that are mandated by state and federal governments or initiated by local leaders. Therefore, staff already are employed and engaged in related activities that can be repurposed as implementation activities with good outcomes. The “cost” is the purposeful effort required on the part of everyone involved to initiate and manage processes to change hearts, minds, and behavior (Blase, Fixsen, Sims, & Ward, 2015).
 - In Study 2 the staff who were selected to be members of implementation teams need to have their current work rapidly repurposed and their current roles promptly reassigned so that team members could learn the new ways of work and devote their time to accomplishing education improvement goals.
- In Study 1 the time required for SISEP staff to teach, coach, and assess implementation capacity at each level exceeded the staffing available in the SISEP group.
 - As a result, the Active Implementation Hub (Ai Hub; <https://implementation.fpg.unc.edu/>) was developed to provide an asynchronous platform for in-state staff to learn the language and basic concepts related to Active Implementation. The Ai Hub is a good fit with ‘just enough, just in time’ teaching and coaching methods used in Study 2 states.
- The State Capacity Assessment (SCA, version 1) was useful for directing attention to key factors involved in capacity development at the state level. However, the data collection method and the number of items were called into question by the state teams who said the assessment took too long, generated too much discussion, and too often was influenced unduly by one or two members of the scoring group.
 - A conceptual review and item analysis conducted by SISEP staff and colleagues produced a more concise and more precise set of items for use in Study 2 states. An in-person facilitation process by a trained administrator and an improved scoring process were established to standardize data collection across states and over time.
- An analysis of the monthly trip reports produced by each SISEP state liaison after each in-state visit in Study 1 led to the development of a State Capacity Development Plan.
 - The State Capacity Development Plan (SCDP) assists and guides SISEP state liaisons in the first 36 months of their work to develop implementation capacity within state education systems (Ward, Jackson, Cusumano, & Fixsen, 2018). Specifically, the State Capacity Development Plan is an operationalized

roadmap of who will do what, how, and by when to achieve the desired outcomes derived from the development of an effective implementation infrastructure to support use of effective innovations as intended.

- The State Capacity Development Plan was used to guide the work in each new state in Study 2. For example, in months 4-6 the Plan asks the SISEP state liaison to help state staff:

- Clearly define roles and responsibilities of STS, document the roles including SMT point of contact or Executive Sponsors (e.g., position description). Include protocols for STS engagement in upper leadership decisions; STS selection criteria and selection process. Develop and use interview protocols that include behavior rehearsals and skill assessment linked to STS position description.

The learning from Study 1 states was integrated into SISEP implementation and scaling methods and applied in Study 2 states. Based on experiences in Study 1, the work in Study 2 states continues to be conducted in improved ways to assure SMT engagement and STS employment in the state education system to quickly establish work at the regional and district levels using new teaching methods. SISEP is learning by doing the work, rapidly examining outcomes (another example of recursive feedback), and promptly revising methods using the Active Implementation Frameworks and associated activities.

Measures

The revised State Capacity Assessment (SCA; Fixsen, Ward, Duda, Horner, & Blase, 2015) addresses critical features of implementation capacity at the state level across three areas: SMT Investment, System Alignment, and Commitment to Regional Implementation. Table 4 provides an overview of the revised SCA subscales and items.

Table 4. State Capacity Assessment and Subscales

SCA (25 Items)	
SMT Investment – 12 items that address roles and functions of the team, executive leadership, and communication of support.	
	Roles and Functions – Examines the composition and meeting processes of the SMT.
	Coordination for Implementation – Identifies the SMT as providing executive leadership in implementation capacity development with needed

SCA (25 Items)	
SMT Investment – 12 items that address roles and functions of the team, executive leadership, and communication of support.	
	resources (e.g., State Transformation Specialists, funding, and access to leadership) to support the work
	Leadership – Includes SMT as providing on-going support for implementation and scaling
System Alignment – 5 items that address development of a State Design Team (SDT) that works to transform legacy systems to effective and efficient practices across a state education system	
	Implementation Guidance Documents – Address transition from a ghost to a host system
	State Design Team – Examine the composition and effective meeting processes in a State Design Team
Commitment to Regional Implementation – 8 items that address delivery of support and resources to Regional Implementation Teams	
	Resources for Implementation Capacity – Allocation of resources to Regional Implementation Teams (RITs) for implementation capacity development
	Support for RIT Functioning – Action planning that is based on review of information and data about RIT functioning

The revised capacity assessment was designed as an “action assessment.” That is, each item is deemed to be essential to system functioning, each item can be scored by participants who are working in the system, each item can be documented, each item is actionable, and the number of items is limited so that assessments can be conducted efficiently twice each year to monitor progress and inform action planning. An assessment is not considered to be complete until an implementation capacity action plan has been developed and reviewed.

In each active scaling state the SCA is administered twice a year by a trained administrator (a SISEP staff member). The respondents who complete the SCA are directly involved in the change processes at the state level (e.g. Deputy Superintendent, Cabinet members, Division Directors, STSs). Scoring of each item is done by consensus. Each item is read by the administrator who then asks for a score. Each respondent simultaneously provides an indication

of his/her score using a card or some other means to indicate a score of 2 (fully in place), 1 (partially in place), or 0 (not yet in place). If there is not 100% agreement then a brief discussion occurs among respondents regarding the rationales for one score or another. The administrator then conducts the scoring routine again. If there is a clear majority, the minority scorers are asked if they can support the majority score (i.e. consensus), if so, the majority score is recorded and the administrator moves to the next item. If there is not consensus, the majority score is recorded and the minority votes are noted for later discussion. The scores are recorded in a national data base on www.sisep.org. When the last question is asked and answered, the data base immediately provides a graph of the results including any past data points for that state. The SCA score is presented as a percentage of the total possible score of 50 (25 items with a maximum score of 2). The total score, subscale scores, and item scores can be reviewed and downloaded. Given this immediate feedback and access to detailed information, within a few days after each SCA is administered the respondents and others prioritize areas that need attention and develop action plans to improve those areas.

An action assessment is an example of a recursive feedback loop (Edmondson & Moingeon, 1998; Gilpin & Murphy, 2008) where a) actions are taken by leaders to promote the development of implementation and scaling capacity, b) assessment data show the areas of current strength and in need of improvement, c) leaders view the data, modify their “view of the problem,” and plan solutions, and d) the assessment data monitor progress and inform further action planning. The key to recursive feedback is to have information about the results of decisions promptly fed back to the decision makers so that poor decisions can be abandoned, assumptions can be revised, and new decisions can be made. Recursive feedback and action planning are the heart of capacity development for systemic change and define the practice-policy communication that is a key to Active Implementation systemic change.

Results

Figure 4 shows the SCA data for five states in Study 2. The relatively rapid development of implementation capacity is evident with each state moving from baseline scores in the 20 – 40% range and reaching the 60% acquisition goal within 24 months. Work in State #6 ended after 26 months with a state election and leadership change. Work began recently in State #10 and the baseline SCA has been administered.

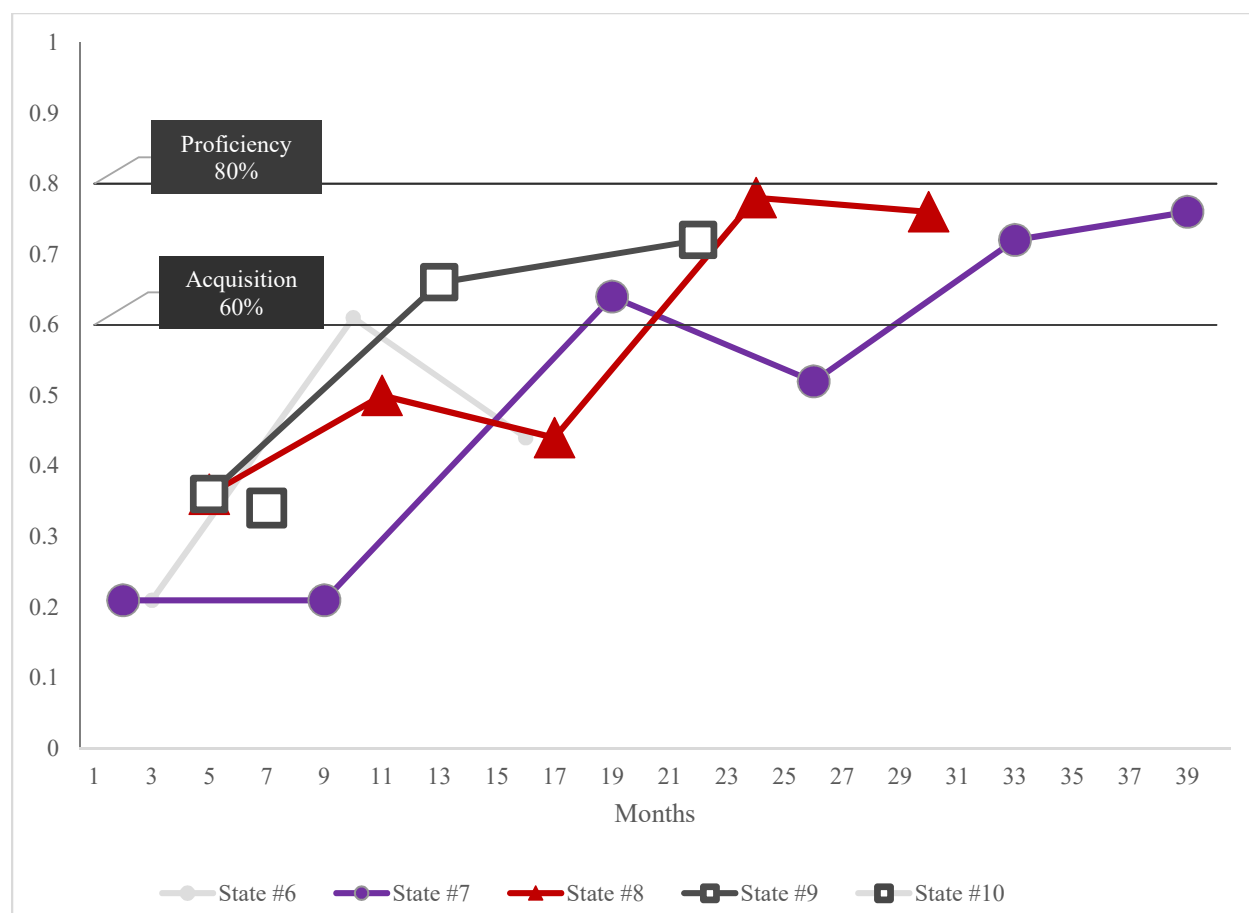


Figure 4. Study 2 State Capacity Assessment (SCA) scores over 39 months.

The SCA subscale scores, shown in Table 5, provide more detailed information. Compared to the other subscale scores, the initial SMT Investment scores are higher at baseline. This is expected since, without leadership investment during the exploration process, implementation and scaling capacity development work would not be happening in that state. Commitment to RIT development is 0% or nearly so in each state at baseline. Traditionally, Regional Education Agencies have not been included as a standard component of state education systems and lack identity and expected functions. Thus, regional capacity development requires extra effort to include regions in thinking and planning at the state level and to negotiate

Table 5. Subscale scores for each administration of the SCA in each active scaling state in Study 2.

Month	SMT Investment	SMT System Alignment	Commitment to RIT	SCA Total
State #6				
3	41%	0%	0%	21%
10	73%	50%	67%	61%
16	54%	33%	33%	44%
State #7				
2	42%	0%	17%	21%
9	41%	0%	50%	21%
19	79%	30%	63%	64%
26	63%	30%	50%	52%
33	88%	40%	69%	72%
39	83%	30%	94%	76%
State #8				
5	54%	40%	6%	36%
11	71%	60%	13%	50%
17	58%	50%	19%	44%
State #9				
5	46%	20%	0%	26%
13	100%	50%	25%	66%
22	88%	40%	69%	72%
State # 10				

7 67% 10% 0% 34%

agreements to reallocate regional staff time and resources required to initiate RITs. Over time, SMT System Alignment is the lowest subscale score in each state and improvement is slow. System alignment requires embedding implementation and scaling best practices in official statements and documents. For example, to be fully in place one SCA item requires that the “State Education Agency (SEA) has written and publicly available documents that describe methods for identifying and supporting effective innovations in education” and another requires that “The SEA has written guidance documents that describe or require providing implementation supports to districts as a primary purpose of regional educational agencies (e.g. ESDs, ECs, Service Co-ops, AEAs, ISDs).” Given the official nature of these requirements, they apply to all regions and districts in a state. SISEP begins working with 2 or 3 regions in a Transformation Zone and with a sample of districts and schools in each district. Thus, it is likely that the system alignment scores will improve more substantially as scaling occurs.

Discussion

The data in Figure 4 are a) the first repeated assessments of state capacity development in education, b) the first to show that purposeful development of implementation capacity is possible in complex state education systems, and c) the first to show purposeful capacity development can be replicated across state departments of education that are unique in terms of history, size, and operations. Each of these findings is significant for education. While improving organization and system capacity for change is seen as critical in education and beyond (Barber & Fullan, 2005; Elmore et al., 2014; Flatten et al., 2011; Marzano, 2010; Padgett, Bekemeier, & Berkowitz, 2005), there are few measures of capacity (Goodman et al., 1993) and little evidence of change with repeated assessments of capacity (McGovern et al., 2007). In one study that employed repeated measures, the McGovern et al. (2007) study included a 35-item measure of capacity that was assessed 3 times over 18 months with scores that improved from slightly less to slightly greater than 3.0 on a 5-point scale. In education systems that have been deemed “intractable” (Sarason, 1996), the data in Figure 4 are the first to reflect systematic change across state education systems and provide hope for a better future for students. The approach to developing capacity seems to meet the purpose of capacity to “accomplish something specific” while engaging in a process to “grow, progress, or improve.” (<http://edglossary.org/capacity>).

The data support the aim of the study to evaluate the viability of the OSEP SISEP implementation informed plan. The changes in SCA scores for states in Study 2 likely are due to the intensive implementation-informed interventions employed by SISEP. The counterfactuals (Handley, Lyles, McCulloch, & Cattamanchi, 2018; U.S. Government Accounting Office, 2009) in education provide added credibility for the findings in Study 2. It is not likely that the

systemic changes reflected in the SCA scores in multiple state education systems occurred by chance.

The baseline data are in the 0-40% range for all 10 states. This is an indicator of a large unmet need for implementation capacity in education systems nationally. A decision by OSEP to explore the use of implementation science in education has resulted in a plan of action to develop implementation capacity, methods to teach implementation science-based methods to existing staff in education systems, and measures to assess the strength of implementation capacity, monitor progress, and inform action planning. The development of capacity in states is informed by the Active Implementation Frameworks that provide a guide for action, methods to defragment and integrate system units, and a focus for improved system functioning.

Bryk (2016) and others (Akin et al., 2013; Harrison & Grantham, 2018; Sarason, 1996) advocate for using the trial and learning methods that are part of improvement science to vastly improve education system functioning and student learning. The merits of this approach are apparent as the lessons learned from five states in Study 1 led to noticeably improved outcomes accomplished in less time for states in Study 2. It is expected that the learning from Study 2 will lead to greater effectiveness and efficiency in states participating in future studies.

In a move away from compliance-focused regulations, OSEP is using implementation science to inform its Results Driven Accountability work that requires states to develop State Systemic Improvement Plans (SSIP) and report annually on progress. OSEP is using the SSIP process to support states' uses of implementation science, best practices, and measures in order to improve outcomes for students with disabilities and for students in general. SISEP is using the knowledge gained from the experience with states to improve supports for other technical assistance centers funded by federal and state education agencies. The goal is to expand implementation and scaling knowledge and skills so that other centers can provide increasingly effective supports for state, regional, and district education systems.

While the data presented in this article provide a basis for optimism, it is likely that current methods to develop capacity in state education systems will continue to evolve and improve. Nevertheless, these are the first data to support the hope that capacity can be developed on purpose in a state system and the capacity-developing methods can be replicated across states. The next steps are to link implementation and scaling capacity with teacher instruction and student outcomes.

Footnotes

¹Preparation of this article was supported in part by a grant from the U.S. Department of Education Office of Special Education Programs #H326K080001 and # H326K120004 and Office of Elementary and Secondary Education #S283B120015. However, the contents do not necessarily represent the policy of the U.S. Department of Education, and endorsement by the Federal Government should not be assumed. We thank Rob Horner and George Sugai who graciously provided their intelligence and support from the beginning. We thank Larry Wexler, Ruth Ryder, and Debra Price-Ellingstad at the Office of Special Education Programs for their foresight, thoughtful support, and willingness to go down an untrodden path in education.

References

- Acosta, J., Chinman, M., Ebener, P., Malone, P. S., Paddock, S., Phillips, A., . . . Slaughter, M. E. (2013). An Intervention to Improve Program Implementation: Findings from a Two-Year Cluster Trial of Assets-Getting to Outcomes. *Implementation Science*, 8(1). doi:10.1186/1748-5908-8-87
- Adelman, H. S., & Taylor, L. (2003). On sustainability of project innovations as systemic change. *Journal of Educational and Psychological Consultation*, 14(1), 1-25.
- Akin, B. A., Bryson, S. A., Testa, M. F., Blase, K. A., McDonald, T., & Melz, H. (2013). Usability testing, initial implementation, and formative evaluation of an evidence-based intervention: Lessons from a demonstration project to reduce long-term foster care. *Evaluation and Program Planning*, 41(0), 19-30. doi:<http://dx.doi.org/10.1016/j.evalprogplan.2013.06.003>
- Barber, M., & Fullan, M. (2005). Tri-level development: Putting systems thinking into action. *Education Weekly*, 24(25), 34-35.
- Ben Charif, A., Zomahoun, H. T. V., LeBlanc, A., Langlois, L., Wolfenden, L., Yoong, S. L., . . . Légaré, F. (2017). Effective strategies for scaling up evidence-based practices in primary care: a systematic review. *Implementation Science*, 12(1), 139. doi:10.1186/s13012-017-0672-y
- Beyer, J. M., & Trice, H. M. (1982). The utilization process: A conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly*, 27(4), 591-622.
- Blanchard, C., Livet, M., Ward, C., Sorge, L., Sorensen, T. D., & McClurg, M. R. (2017). The Active Implementation Frameworks: A roadmap for advancing implementation of Comprehensive Medication Management in Primary care. *Research in Social and Administrative Pharmacy*, 13(5), 922–929. doi:10.1016/j.sapharm.2017.05.006
- Blase, K. A., Fixsen, D. L., Sims, B. J., & Ward, C. S. (2015). *Implementation science: Changing hearts, minds, behavior, and systems to improve educational outcomes*. Retrieved from University of North Carolina at Chapel Hill: National Implementation Research Network: <http://www.winginstitute.org/News-And-Events/Summits/2014-Summit-Summary>
- Brachthausen, C. (2011). Explaining global governance – a complexity perspective. *Cambridge Review of International Affairs*, 24(2), 221-224.
- Bryk, A. S. (2016). Accelerating How We Learn to Improve. *Educational Researcher*, 44(9), 467-477. doi:10.3102/0013189X15621543
- Chaple, M., & Sacks, S. (2016). The Impact of Technical Assistance and Implementation Support on Program Capacity to Deliver Integrated Services. *The Journal of Behavioral Health Services & Research*, 43(1), 3-17. doi:10.1007/s11414-014-9419-6

Chinman, M., Hunter, S. B., Ebener, P., Paddock, S. M., Stillman, L., Imm, P., & Wandersman, A. (2008). The getting to outcomes demonstration and evaluation: An illustration of the prevention support system. *American Journal of Community Psychology*, 41, 206-224.

Conklin, J. (2005). Wicked problems and social complexity. In J. Conklin (Ed.), *Dialogue mapping: Building shared understanding of wicked problems* (pp. 1-25). New York: Wiley.

Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan*, 76(8), 642-645.

Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., Graczewski, C., . . . Giffin, J. (2017). *School Improvement Grants: Implementation and Effectiveness (NCEE 2017-4013)*. Retrieved from Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education: <http://ies.ed.gov/ncee>.

Edmondson, A., & Moingeon, B. (1998). From Organizational Learning to the Learning Organization. *Management Learning*, 29(1), 5-20. doi:10.1177/1350507698291001

Elmore, R. F. (2002). *Bridging the gap between standards and achievement: The imperative for professional development in education*. Retrieved from Washington, DC: The Albert Shanker Institute:

Elmore, R. F., Forman, M. L., Stosich, E. L., & Bocala, C. (2014). *The Internal Coherence Assessment Protocol & Developmental Framework: Building the Organizational Capacity for Instructional Improvement in Schools*. Retrieved from Washington, DC: Strategic Education Research Partnership Institute:

Epstein, D., & Klerman, J. A. (2013). When is a program ready for rigorous impact evaluation? The role of a falsifiable logic model. *Evaluation Review*, 36, 375-401. doi:10.1177/0193841X12474275

Firestein, S. (2016). *Failure: Why science is so successful*. Oxford University Press.

Fixsen, D. L., Blase, K., Naoom, S., Metz, A., Louison, L., & Ward, C. (2015). *Implementation Drivers: Assessing Best Practices*. Retrieved from National Implementation Research Network, University of North Carolina at Chapel Hill:

Fixsen, D. L., Blase, K. A., Duda, M., & Horner, R. (2009). *State capacity assessment*. Retrieved from Chapel Hill, NC:

Fixsen, D. L., Blase, K. A., Duda, M., Naoom, S., & Van Dyke, M. (2010). Implementation of evidence-based treatments for children and adolescents: Research findings and their implications for the future. In J. Weisz & A. Kazdin (Eds.), *Implementation and dissemination: Extending*

treatments to new populations and new settings (2nd ed., pp. 435-450). New York: Guilford Press.

Fixsen, D. L., Blase, K. A., & Fixsen, A. A. M. (2017). Scaling effective innovations. *Criminology & Public Policy*, 16(2), 487-499. doi:10.1111/1745-9133.12288

Fixsen, D. L., Blase, K. A., Horner, R., & Sugai, G. (2009). *Intensive technical assistance*. Retrieved from Chapel Hill, NC:

Fixsen, D. L., Blase, K. A., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based programs. *Exceptional Children (Special Issue)*, 79(2), 213-230.

Fixsen, D. L., Blase, K. A., Metz, A., & Van Dyke, M. (2015). Implementation science. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd ed., Vol. 11, pp. 695-702). Oxford: Elsevier, Ltd.

Fixsen, D. L., Blase, K. A., & Van Dyke, M. (2012). *From ghost systems to host systems via transformation zones*. Retrieved from Washington, DC: U.S. Department of Education Office of Vocational and Adult Education: <http://www.lincs.ed.gov>

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, National Implementation Research Network.

Fixsen, D. L., Ward, C., Ryan Jackson, K., & Chaparro, E. (2015). *Observation tool for instructional supports and systems (OTISS)*. Retrieved from National Implementation Research Network, University of North Carolina:

Fixsen, D. L., Ward, C. S., Duda, M. A., Horner, R., & Blase, K. A. (2015). *State Capacity Assessment (SCA) for Scaling Up Evidence-based Practices* (v. 25.2). Retrieved from Chapel Hill, NC: National Implementation Research Network, State Implementation and Scaling up of Evidence Based Practices Center, University of North Carolina at Chapel Hill:

Flatten, T. C., Engelen, A., Zahra, S. A., & Brettel, M. (2011). A measure of absorptive capacity: Scale development and validation. *European Management Journal*, 29(2), 98-116. doi:<https://doi.org/10.1016/j.emj.2010.11.002>

Fox, J., & Gershman, J. (2000). The World Bank and Social Capital: Lessons from Ten Rural Development Projects in the Philippines and Mexico. *Policy Sciences*, 33(3-4), 399-419.

Freeman, R., Perrin, N., Irvin, L., Vincent, C., Newcomer, L., & Moore, M. (2010). *Positive behavior support across the lifespan: Expanding the concept of statewide planning for large scale organizational systems change*. Retrieved from Lawrence, KS:

Frick, T., Elder, M., Hebb, C., Wang, Y., & Yoon, S. (2006). *Adaptive usability evaluation of complex web sites: How many tasks?* Retrieved from Bloomington, IN: School of Education

Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press. .

Gilpin, D. R., & Murphy, P. J. (2008). *Crisis management in a complex world*. New York: Oxford University Press.

Goodman, R. M., McLeroy, K. R., Steckler, A. B., & Hoyle, R. H. (1993). Development of Level of Institutionalization Scales for Health Promotion Programs. *Health Education Quarterly*, 20(2), 161-178. doi:10.1177/109019819302000208

Gotham, H. J., Brown, J. L., Comaty, J. E., McGovern, M. P., & Claus, R. E. (2013). Assessing the Co-occurring Capability of Mental Health Treatment Programs: the Dual Diagnosis Capability in Mental Health Treatment (DDCMHT) Index. *The Journal of Behavioral Health Services & Research*, 40(2), 234-241. doi:10.1007/s11414-012-9317-8

Hall, G., & Hord, S. M. (1987). *Change in schools: Facilitating the process*. Albany, NY: SUNY Press.

Handley, M. A., Lyles, C. R., McCulloch, C., & Cattamanchi, A. (2018). Selecting and Improving Quasi-Experimental Designs in Effectiveness and Implementation Research. *Annual Review of Public Health*, 39(1), 5-25. doi:10.1146/annurev-publhealth-040617-014128

Harrison, M. I., & Grantham, S. (2018). Learning from implementation setbacks: Identifying and responding to contextual challenges. *Learning Health Systems*, 0(0), e10068. doi:doi:10.1002/lrh2.10068

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Hord, S. M. (1992). *Facilitative leadership: The imperative for change*. Austin, TX: Southwest Educational Development Laboratory.

Horner, R. H., Sugai, G., & Fixsen, D. L. (2017). Implementing effective educational practices at scales of social importance. *Clinical Child and Family Psychology Review*, 20(1), 25-35. doi:10.1007/s10567-017-0224-7

Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B. (2004). The School-wide Evaluation Tool (SET): A Research Instrument for Assessing School-wide Positive Behavior Support. *Journal of Positive Behavior Interventions*, 6(1), 3-12.

Jiménez-Barrionuevo, M. M., García-Morales, V. J., & Molina, L. M. (2011). Validation of an instrument to measure absorptive capacity. *Technovation*, 31(5), 190-202.

doi:<https://doi.org/10.1016/j.technovation.2010.12.002>

Kluger, J. (2008). *Simplexity: Why simple things become complex (and how complex things can be made simple)*. New York: Hyperion.

Lee, N., & Cameron, J. (2009). Differences in self and independent ratings on an organisational dual diagnosis capacity measured. *Drug and Alcohol Review*, 28, 682-684. doi:10.1111/j.1465-3362.2009.00116.x

Manna, P. (2008). *Federal aid to elementary and secondary education: Premises, effects, and major lessons learned*. Williamsburg, VA: College of William and Mary, Department of Government and the Thomas Jefferson Program in Public Policy.

Marzano, R. (2010). *Teacher scales for reflective practice: Applying the art and science of teaching*. Englewood, Colorado: Marzano Research Laboratory: Englewood, Colorado: Marzano Research Laboratory.

Marzano, R., Waters, T., & McNulty, B. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

McGovern, M. P., Matzkin, A. L., & Giard, J. (2007). Assessing the Dual Diagnosis Capability of Addiction Treatment Services: The Dual Diagnosis Capability in Addiction Treatment (DDCAT) Index. *Journal of dual diagnosis*, 3(2), 111-123. doi:10.1300/J374v03n02_13

McIntosh, K., Mercer, S. H., Nese, R. N. T., & Ghemraoui, A. (2016). Identifying and Predicting Distinct Patterns of Implementation in a School-Wide Behavior Support Framework. *Prevention Science*, 17(8), 992-1001. doi:10.1007/s11121-016-0700-1

Moll, L. C. (1990). Introduction to Vygotsky and Education: Instructional implications of sociohistorical psychology. In L. C. Moll (Ed.), *Vygotsky and education: Instructional implications of sociohistorical psychology* (pp. 1-30). Cambridge: Cambridge University Press.

Morgan, G. (1997). *Images of organization* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Morgan, G., & Ramirez, R. (1983). Action learning: A holographic metaphor for guiding social change. *Human Relations*, 37(1), 1-28.

National Center for Education Statistics. (2013). *The Nation's Report Card: Trends in Academic Progress 2012*. Retrieved from Institute of Education Sciences, U.S. Department of Education: <http://nces.ed.gov/nationsreportcard/subject/publications/main2012/pdf/2013456.pdf>

National Commission on Excellence in Education. (1983). A Nation at Risk: The Imperative for Educational Reform. *The Elementary School Journal*, 84(2), 113-130. doi:10.1086/461348

- Nielsen, J. (2005). Usability for the masses. *Journal of Usability Studies*, 1(1), 2-3.
- Padgett, S. M., Bekemeier, B., & Berkowitz, B. (2005). Building sustainable public health: Systems change at the state level. *Journal of Public Health Management and Practice*, 11(2), 109-115.
- Petroski, H. (1982). *To engineer is human: The role of failure in successful design* (Vintage ed.). New York: Random House.
- Resnick, L. B. (2010). Nested system for the thinking curriculum. *Educational Researcher*, 39(3), 183-197.
- Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155-169.
- Sabatier, P. A. (1986). Top-Down and Bottom-Up Approaches to Implementation Research: a Critical Analysis and Suggested Synthesis. *Journal of Public Policy*, 6(01), 21-48.
doi:doi:10.1017/S0143814X00003846
- Saetren, H. (2014). Implementing the third generation research paradigm in policy implementation research: An empirical assessment. *Public Policy and Administration*, 29(2), 84-105. doi:10.1177/0952076713513487
- Sarason, S. B. (1996). *Revisiting "The culture of the school and the problem of change"*. New York: Teachers College Press.
- Senge, P. M. (2006). *The fifth discipline: The art and practice of the fifth discipline*. New York: Doubleday.
- Singleton, R., Straits, B. C., Straits, M. M., & McAllister, R. J. (1988). *Approaches to social research*. New York: Oxford University Press.
- St. Martin, K., Ward, C., Harms, A., Russell, C., & Fixsen, D. L. (2015). *Regional capacity assessment (RCA) for scaling up implementation capacity*. Retrieved from National Implementation Research Network, University of North Carolina:
- Svensson, J., Tomson, K., & Rindzeviciute, E. (2017). Policy change as institutional work: introducing cultural and creative industries into cultural policy. *Qualitative Research in Organizations and Management: An International Journal*, 12(2). doi:doi:10.1108/QROM-05-2016-1380
- Tyack, D., & Cuban, C. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.

U.S. Department of Education. (2009). The federal role in education. Retrieved from <http://www.ed.gov/about/overview/fed/role.html>

U.S. Government Accounting Office. (2009). Program evaluation: A variety of rigorous methods can help identify effective interventions. Retrieved from <http://gao.gov/products/GAO-10-30>

Ulrich, W. M. (2002). *Legacy systems: Transformation strategies*. Upper Saddle River, NJ: Prentice Hall PTR.

Vernez, G., Karam, R., Mariano, L. T., & DeMartini, C. (2006). *Evaluating comprehensive school reform models at scale: Focus on implementation*. Retrieved from Santa Monica, CA: <http://www.rand.org/>

Ward, C., Jackson, K. R., Cusumano, D., & Fixsen, D. (2018). *State Capacity Development Plan*. Retrieved from University of North Carolina at Chapel Hill: National Implementation Research Network:

Ward, C., St. Martin, K., Horner, R., Duda, M., Ingram-West, K., Tedesco, M., . . . Chaparro, E. (2015). *District Capacity Assessment*. Retrieved from National Implementation Research Network: University of North Carolina at Chapel Hill:

Zucker, L. G. (1987). Institutional theories of organizations. *Annual Review of Sociology*, 13, 443-464.