Published in final edited form as:

Child Dev. 2017 September ; 88(5): 1435–1446. doi:10.1111/cdev.12904.

From Innovation to Impact at Scale: Lessons Learned from a Cluster of Research-Community Partnerships

Holly S. Schindler,

University of Washington, College of Education, Miller Hall Box 353600, Seattle, WA 98105

Philip A. Fisher, and

University of Oregon

Jack P. Shonkoff

Harvard University

Abstract

This paper presents a description of how an interdisciplinary network of academic researchers, community-based programs, parents, and state agencies have joined together to design, test, and scale a suite of innovative intervention strategies rooted in new knowledge about the biology of adversity. Through a process of co-creation, collective pilot-testing, and the support of a measurement and evaluation hub, the Washington State Innovation Cluster is using rapid cycle, iterative learning to elucidate differential impacts of interventions designed to build child and caregiver capacities and address the developmental consequences of socioeconomic disadvantage. Key characteristics of the Innovation Cluster model are described and an example is presented of a video-coaching intervention that has been implemented, adapted, and evaluated through this distinctive, collaborative process.

Human development proceeds at a rapid pace during the first five years of life, making young children particularly sensitive to both adverse and positive experiences, as well as responsive to interventions (Fox, Levitt, & Nelson, 2010; Shonkoff & Phillips, 2000). Decades of program evaluation research have demonstrated that interventions during early childhood can have a range of positive impacts (Camilli, Vargas, Ryan, & Barnett, 2010; Gorey, 2001), especially for those experiencing socioeconomic disadvantages (Gormley, Phillips, & Gayer, 2008; Yoshikawa et al., 2013), but the magnitude of those impacts has been variable and typically modest (Duncan & Magnuson, 2013). Moreover, no major program innovations in recent years have produced breakthrough outcomes or achieved population-level effects at scale. Stated simply, science and policy have not mitigated the fact that large numbers of young children continue to face the burdens and consequences of significant adversity.

Advances in neuroscience and the biology of adversity now offer unprecedented new insights about what young children need for healthy development (Bruce et al., 2013; Gunnar & Fisher, 2006; Shonkoff, 2012). This science underscores the importance of

mutually contingent, “serve and return” interactions between young children and adults for strengthening developing brain architecture (National Scientific Council on the Developing Child, 2004, 2012). There is also growing evidence that the foundations of self-regulation and executive function skills—which are essential for academic success, mental health, economic self-sufficiency, effective parenting, and resilience in the face of adversity—can be undermined when unresponsive caregiving fails to provide buffering protection against significant stress in the early childhood years (Center on the Developing Child, 2011; National Scientific Council on the Developing Child, 2005/2014; Shonkoff, et. al, 2012).

This rapidly advancing science calls for a new early childhood agenda that builds on current investments in quality improvement and system building and seeks new models and methods in the quest for greater impacts. To this end, there is a need for enhanced theories of change and more effective strategies that move beyond the general question of “what works?” and seek a more nuanced understanding of what works (and what doesn’t) *for whom and why, and in what contexts* (Shonkoff & Fisher, 2013). This paper presents a description of how an interdisciplinary network of researchers, community-based practitioners, parents, and state policymakers has responded to this challenge through a unique approach to co-designing, testing, and scaling a suite of innovative, science-based, intervention strategies to achieve breakthrough outcomes for young children facing adversity. This network, known as the Washington State Innovation Cluster, is embedded in the Frontiers of Innovation (FOI), a rapidly-growing learning community created by the Center on the Developing Child at Harvard University to catalyze experimentation, risk-taking, and learning from failure.

FOI is driven by a translational science model (TSM) that uses time-tested tools in a new way. This approach to program design and evaluation requires specification of precise and measureable intervention targets, mediators, and outcomes. Its rapid cycle and iterative nature embraces lessons learned from experimental medicine about how intervention trials can serve as tests of underlying theoretical models and thereby provide clarity about whether changes in specific target domains are causally related to changes in pre-determined outcomes. FOI theories of change also place a high priority on identifying person- and place-based moderators of intervention efficacy. This allows new interventions, as they undergo iterations of evaluation, to be matched more closely to the needs of specified subgroups in different contexts. This paper summarizes the key elements of the Washington Innovation Cluster that have been critical to its success to date, and describes an illustrative intervention strategy that has been developed and piloted within this model.

Washington State Innovation Cluster

The notion of an innovation cluster was created in 2011 by a cadre of state-level policy makers who coined the term “one science approach” to coordinate investments in young children across agencies. Energized by the Harvard Center’s formulation of an integrated science of learning, behavior, and health (Center on the Developing Child, 2010), these government leaders convened an initial group of community-based service providers and university-based researchers to explore new ways of strengthening the capabilities of young children and their caregivers, particularly those at the highest risk of experiencing toxic stress. Over the ensuing five years, members of this evolving network have engaged in a

design and evaluation process characterized by three interdependent elements that differ from conventional intervention research. These include: (1) co-creation of new strategies by researchers, practitioners, and parents; (2) collective pilot testing of diverse approaches linked by a common goal; and (3) a research and evaluation hub that facilitates rapid cycle learning and continuous feedback, with an eye towards scaling. Current research institutions, intervention strategies, and innovating sites are shown in Figure 1 and details of the Cluster's synergistic strategy are described below.

Co-Creation

Despite the widely acknowledged value of research that extends beyond highly controlled, laboratory settings (Bronfenbrenner, 1977), authentic researcher-practitioner partnerships in community sites are not common in the early childhood arena. Though they have a more robust history in other fields, such as K-12 education (Tikunoff & Ward, 1983; Wagner, 1997), the most typical form of collaboration involves investigators who seek permission from community-based programs to study issues of academic interest. The Washington State Innovation Cluster, in contrast, began with interdisciplinary project teams of researchers, practitioners, and service recipients who jointly formulated research questions to address important challenges or unmet needs, share data collection, and reflect on their findings. These teams design and test strategies within existing programs (e.g., early childhood education, home visiting) and seek ongoing input from a working group of program directors, managers, and senior advisors from multiple agencies. This co-creation process assures that each new intervention strategy is relevant in a real-world context and has scaling potential within currently functioning systems.

Collective Pilot Testing

The benefits of researcher-practitioner-parent collaboration in testing new intervention models are augmented through multiple opportunities for connecting with other project teams in the Cluster. These teams pursue shared goals in building both child and caregiver capabilities and addressing the developmental consequences of early life stress. Like Networked Improvement Communities that connect school districts around a common problem (Coburn, Penuel, & Geil, 2013), the Washington State Innovation Cluster teams collectively pilot test multiple strategies in diverse contexts to better understand what works (or doesn't work) for whom and under what conditions. Currently, six strategies are being tested, some of which have undergone pilots in multiple sites (See Figure 1) and have been adapted for multiple contexts (e.g., classrooms, family child care settings, home visits). Through this process, the Cluster is developing a suite of intervention strategies (rather than a single program) and identifying combinations of strategies that together can deliver the largest benefits for different subgroups of children and families. Although each pilot is initially small, its modest size enables fast-cycle testing, which promotes greater understanding of its active ingredients and real-time sharing across the Cluster.

A Research and Evaluation Hub

A productive, shared learning process requires that each strategy in the Cluster generates the following: (a) detailed intervention materials; (b) an explicit theory of change (TOC); and (c) an evaluation plan that is closely tied to the TOC. Most researchers and practitioners had

some, but not extensive, prior experience with these components. As the Cluster evolved, the level of performance needed for fast-cycle learning and the construction of aggregate knowledge across projects demanded more than the typical expectations for program development and testing. In that context, the need for a research and evaluation hub to provide tools and consultation for each project team throughout the design and testing cycle became clear. Thus, each team is matched with a highly skilled contact person from the hub who provides most of the consultation. Next, an information exchange and follow-up are held in which the project team shares its current thinking and an action plan is created. Technical assistance is then provided by the hub point person, as needed, to ensure that the intervention materials, theories of change, and evaluation plans are highly specific and well-aligned. The rationale for each component is described below.

Intervention materials—Current standards of evidence for research in prevention science indicate that most interventions are designed to work well in tightly-controlled, demonstration projects but may not be designed from the outset to be taken to scale in real-world settings (Gottfredson et al., 2015). Thus, Gottfredson and colleagues suggest that the next generation of preventive interventions should, from the start, “be described at a level that would allow others to implement/replicate” them. The Washington Innovation Cluster illustrates this approach by demanding a clearly defined intervention as a critical prerequisite for both feasibility testing and scalability. To this end, the Cluster’s research and evaluation hub provides support around the development of intervention materials (e.g., parent handouts and practitioner guides) that provide detailed information on the program’s theoretical framework, primary population, intervention targets, and service setting. Program materials are also required to detail *how* the service provider will promote change in the intervention targets, including what she or he will say or do in each session. These materials serve two important purposes: (1) to act as guides for service providers who implement the intervention; and (2) to provide a means for assessing fidelity when the intervention is replicated or scaled.

Theory of change—A TOC is a detailed set of testable beliefs about specific changes that are expected to result from an intervention focused on explicit, target variables. Identifying the hypothesized causal mechanisms (i.e., mediators) underlying intervention effects allows members of the Cluster to unpack the “black box” of their strategies. As noted by MacKinnon and Luecken (2008), analysis of mediation is necessary when a field reaches a juncture at which it becomes critical to create more effective and efficient approaches. The TOC for each intervention strategy in the Cluster also articulates hypotheses about moderators of intervention effectiveness, including contextual and individual variables that specify which participants are expected to benefit more from an intervention and which participants are expected to benefit less or not at all. Although moderated intervention effects have historically been viewed as disappointing, based on the hope that a single program will affect all participants, the Cluster views moderated impacts that are theoretically articulated prior to field testing as important opportunities to iteratively create more efficacious strategies that reach more individuals. Moreover, moderators “represent the parameters that enable investigators to make precise predictions regarding the processes that guide behavior and, in particular, to specify the conditions under which these processes

operate” (Rothman, 2013, p. 190). Stated simply, the TOC serves as a tool to help researcher-practitioner teams replace the question “does it work?” with queries about how, why, and for whom does an intervention work or not. Although this alternative approach has been advocated by prevention scientists for many years (e.g., Gottfriedson et al., 2015; MacKinnon & Luecken, 2008), its implementation has been limited.

The research and evaluation hub in the Washington Cluster supports each researcher-practitioner-parent team in creating its own TOC through written materials, workshops, and consultation as needed. Each TOC is based on a common template (see Figure 2) that focuses on testing causal mechanisms (i.e., mediators) related to building adult capabilities (e.g., executive function skills) as well as moderators (e.g., adverse experiences). As expected and unexpected findings related to individual TOCs are shared the collective knowledge gained from testing multiple TOCs reveals opportunities for improving existing strategies or hybridizing two or more promising approaches. Most importantly, the benefits of this model depend on the design and implementation of evaluation plans linked to clearly defined theories of change.

Evaluation plan—Each Cluster project team working on an intervention strategy develops a rigorous evaluation plan that enables the testing of precise hypotheses derived from its TOC. To support shared learning across the Cluster, a set of common metrics and measures that are recommended by the research and evaluation hub are used when appropriate, and all data are entered into a shared database. In addition to analyzing individual TOCs, this allows the Cluster to answer common research questions that span multiple intervention strategies.

Another unique feature of the Cluster’s evaluation methodology is its strategic utilization of micro-trials (stemming in part from methods proposed by Howe et al., 2010). In contrast to a direct linear route from efficacy trials to experimental effectiveness trials and then on to dissemination, each intervention strategy is evaluated initially through a series of low-cost, pilot tests with a small number of children and caregivers. The purpose of these early micro-trials is to establish evidence of feasibility and to conduct a preliminary evaluation of the TOC. Subsequent micro-trials use both positive and negative findings to refine the intervention strategy’s intervention materials, TOC, and evaluation plan. The assumption underlying this approach is that an initial series of micro-trials iterating specific intervention components is likely to yield more valuable information than one large-scale trial. This process does not preclude the use of randomized controlled trials or effectiveness trials at a later point, but simply argues for using the approach that is best-matched to the questions at hand. For example, as intervention strategies are refined, they may lead to randomized clinical trials to answer questions about longitudinal impacts. This process of fast-cycle feedback continues until scalable findings are achieved.

Case Example: Filming Interactions to Nurture Development (FIND)

One promising strategy that is being implemented and evaluated within the Washington Innovation Cluster is a strength-based video coaching program called Filming Interactions to Nurture Development (FIND; for details see Fisher et al. in press). Though originally developed for mothers, FIND is also being tested in the Cluster with low-income fathers

(FIND-F). Unlike typical didactic fathering groups, FIND-F is delivered through individual home visits, targets responsive father-child interactions, and uses video to identify parenting strengths. Since 2012, the federal government has invested nearly \$1.5 billion in home visiting programs through Maternal, Infant, and Early Childhood Home Visiting (MIECHV), which funds services for over 100,000 families across the United States (U.S. Department of Health and Human Services, 2015). At the same time, the government is also investing millions of new dollars into responsible fatherhood programs. However, to date, there has been minimal coordination between these two efforts. FIND-F's brief, flexible, strength-based model may offer a strategy for leveraging existing home visiting services and responsible fathering programs to achieve greater impacts. This section describes how key Cluster features have supported its initial development and testing.

The Co-Creation Process

The FIND-F Cluster project team included researchers from the University of Washington, FIND program developers at the University of Oregon, and practitioners and fathers from Children's Home Society of Washington, a provider of comprehensive services for young children and families. The FIND-F model was developed in response to the challenges cited by program staff in recruiting, retaining, and supporting low-income fathers through two of their home visiting programs in which 70% of enrolled families had a resident father but few were participating regularly.

During the development phase, researchers met weekly with two home visitors and a program supervisor. Early meetings focused on relationship building and discussions about how best to engage families and recruit fathers. Early in this period, the project team determined that the decision making process would benefit from the voices of fathers and additional practitioners. Consequently, over the course of several months, the team conducted semi-structured interviews with 10 fathers and 5 home visitors using an interview guide in which a number of constructs were designated for discussion but the tempo and order of the conversation flowed from the respondents' lead (Patton, 2002). This interview guide included questions about fathers' perceived needs and resources, daily routines, best practices for informing them of program availability, best practices for making the program attractive to men, potential barriers to program involvement, and reactions to the FIND-F format and content. A similar approach using focus groups has been found to be successful in informing intervention development for other hard-to-reach populations (Julion, Breitenstein, & Waddell, 2012; Lengua, et al., 1992).

Analyses of the interview data revealed a number of critical insights about how to make the program feasible and effective for fathers. These included the need to: (1) reduce the number of coaching sessions from 10 to 6; (2) offer services at different times of the day, including evenings and weekends; (3) ask fathers about their comfort level with male versus female home visitors (and provide male home visitors when needed); and (4) focus filming and coaching sessions on fathers, not couples. Subsequent meetings between the researchers and practitioners focused on co-creating intervention materials, articulating FIND-F's TOC, and developing an evaluation plan. Multiple sources of information were used, including interview data and both practitioner and researcher expertise. For example, practitioners

provided key insights regarding the validity, feasibility, and value to their day-to-day work of each measurement tool. In one case, the team jointly agreed it would be important to use a standard observational measure of fathers' parenting skills, and the University of Washington team then investigated several such measures and presented them to the group. The instrument that the team ultimately selected was chosen in part because Children's Home Society of Washington saw its potential utility as a practical and versatile tool for their home visitors beyond the research study. The team also received valuable input from the larger Innovation Cluster on how to simplify its TOC, as well as guidance from the research and evaluation hub on selecting measures for common constructs (e.g., demographic variables and child outcomes) and assessing the strengths and limitations of different evaluation plans.

Although the co-creation process was successful, it also faced challenges. One critical lesson learned was the importance of considering the time demands on service providers and the need to work with the relevant agencies to provide additional compensation to participating practitioners or to create space within their caseloads to take on this additional commitment. Removing barriers to practitioner participation is especially important for recruiting individuals who reflect the diverse characteristics of the service recipients. Securing strong support within the management of the agency is also critical. For example, one of the practitioners was initially hesitant to give up part of his caseload to work on the project unless he received assurance from his supervisors that he could regain his caseload after the project ended. Ultimately, he participated fully in the process because he trusted the commitment of his manager, who has been part of the Washington Innovation Cluster since its inception.

FIND-F's Theory of Change

Though most commonly conceptualized in terms of maternal responsiveness, healthy serve-and-return interactions between fathers and their children are increasingly recognized as important influences on child development (e.g., Cabrera, Shannon, & Tamis-LeMonda, 2007). In spite of these findings, however, only a small subset of interventions for low-income fathers has focused on their parenting skills (see Cowan, Cowan, & Knox, 2010, for a review) and none has focused specifically on their responsiveness to young children. FIND-F addresses this gap by directly targeting fathers' responsive parenting through in-home coaching based on films of their own, naturally occurring interactions. More specifically, a FIND-F coach visits the father and child in the father's home once a week for 6 weeks. In each session, the coach takes a short film of the father and child engaging in everyday activities, such as playing with toys or eating a meal. Next, the film is analyzed by trained editors and clips are selected to highlight strengths in the father-child interaction. These clips are then reviewed with the father the following week. The aim of this intervention is to reinforce the types of positive interactions that are hypothesized to lead to increased behavioral and psychological engagement and decreased parenting stress for fathers as well as fewer behavior problems in the children. Though not measured in this initial pilot due to cost and feasibility constraints (i.e., many children in the sample were too young for currently developed executive function measures), we believe these benefits are likely to be facilitated by improvements in father and child executive function skills, such as attention

and impulse control. For example, fathers are coached to wait for their children's cues during serve and return interactions, a skill that requires fathers to practice inhibition while withholding a prepotent response (for additional details see Fisher et al., in press). We also hypothesized that fathers who experienced high levels of adversity in their own childhood would benefit the most from FIND-F's strength-based, skill-building approach. See Figure 3 for a graphical representation of FIND-F's theory of change.

Methods

A small pretest-posttest micro-trial was employed in the initial testing to establish evidence of feasibility and conduct an initial evaluation of the underlying TOC. Although micro-trials limit the ability to draw strong conclusions, when conducted rigorously they can provide valuable empirical information at a relatively low cost.

Sample—The recruited sample for the initial pretest-posttest micro-trial included 15 fathers from two home visiting programs serving low-income families. Each participating father had to speak English or Spanish fluently and have a child between 6 and 36 months of age.

Measures—Feasibility was measured by the number of recruited fathers who completed FIND-F, the number of weeks it took to reach completion, and the percentage of sessions held on evenings and weekends. The intervention target, responsive parenting skills, was observed via videotaped, father-child interactions and coded using the PICCOLO-D protocol (Parenting Interactions with Children Checklist of Observations Linked to Outcomes- Dad version) (Anderson, Roggman, Innocenti, & Cook, 2013). Fathers also reported on a number of parent and child outcomes. Parenting stress was measured using the Parenting Stress Index Short Form (PSI-SF) (Loyd & Abidin, 1985), behavioral father involvement was measured by the Who Does What questionnaire (WDW; Cowan & Cowan, 1990), and psychological father involvement was measured through the PIE (Cowan & Cowan, 1991). Children's problem behaviors were measured using an infant/toddler version of the Parent Daily Report (PDR) (Chamberlain & Reid, 1987). To measure the moderator, childhood adversity, fathers were asked about whether they had experienced 15 different adverse childhood experiences (ACEs). Fathers with three or more ACEs were considered to have experienced high levels of adversity. See Technical Appendix A for more detailed descriptions of measures.

Data analysis strategy—Changes from pretest to posttest on the intervention target, father outcomes, and child outcomes were assessed through a series of paired samples t-tests. We report and interpret effect sizes in addition to statistical significance. T-tests and effect sizes were also calculated separately for fathers with low and high ACEs to explore potential moderation patterns. Given the modest sample size and the limitations regarding statistical power in this micro-trial, effect sizes are likely to be more meaningful than significance level.

Results

Feasibility—Of the 15 recruited fathers, 12 (80%) completed all six sessions of FIND-F. The average length of time it took for fathers to complete the six sessions was 6.10 weeks ($SD=1.78$). Thirty-eight percent of the sessions were held in the evenings, and 24% on weekends.

Descriptive statistics—Fathers' ages ranged from 22 to 56 years ($M=34.75$). Races and ethnicities of fathers included Hispanic ($n=7$), African American ($n=2$), Caucasian ($n=2$), and mixed Caucasian and Native American ($n=1$). Two fathers had less than a high school education, seven had a high school diploma or GED, and three completed education beyond high school. All fathers were living with a partner; seven (58.33%) were married. All fathers were working either full-time ($n=9$) or part-time ($n=3$), with 80% reporting work on weekends and 50% reporting irregular work schedules. The number of children in the household ranged from one to seven ($M=3$). Children's ages ranged from 9 to 35 months ($M=22.67$); four (33.33%) were female and eight (66.67%) were male. Five fathers (42%) reported high ACE scores (>3).

Intervention targets—The remaining results are detailed in Table 1. Fathers' responsive parenting skills significantly increased from 31.30 ($SD=6.92$) to 34.90 ($SD=6.92$). This change is equivalent to moving from “barely” engaging in a particular skill to “clearly” engaging in a particular skill on three to four out of 21 items ($d=.90$, $p=.02$). The pattern of effects was similar for fathers with low ($d=.99$, $p=.06$) and high ($d=.69$, $p=.26$) ACE scores.

Father outcomes—Fathers' parenting stress scores significantly decreased from 73.27 ($SD=15.39$) to 63.08 ($SD=13.95$). This change corresponds to moving from the 50th percentile on parenting stress to the 30th percentile ($d=-1.31$, $p=.001$). Changes were similar for fathers with low ($d=-1.35$, $p=.01$), and high ($d=-1.39$, $p=.04$) ACE scores.

Fathers' mean level of behavioral involvement increased from 3.69 ($SD=.99$) to 3.89 ($SD=1.14$) in the direction of more equal division of caretaking tasks. This change was statistically insignificant but represented a medium effect size ($d=.44$, $p=.17$). In conducting analyses separately for fathers with high and low ACE scores, the growth from baseline to post-test seems to be driven by fathers with high scores. The effect size for fathers with low ACEs was small and statistically insignificant ($d=.07$, $p=.85$) while the effect size for fathers with high ACEs was large and marginally significant ($d=.96$, $p=.09$).

Fathers' baseline mean psychological involvement was 125.58 ($SD=64.50$), suggesting that, on average, fathers attributed 125.58 degrees of a circular “pie” representing their identities for the fathering role. After 6 weeks of FIND-F, this increased to 143.92 ($SD=64.54$), a change that was statistically insignificant but represented a medium effect size ($d=.50$). Again, the change from baseline to posttest seemed to be driven by fathers who had experienced the most childhood adversity. The change for fathers with low ACEs was small, in the negative direction, and statistically insignificant ($d=-.08$, $p=.84$) while the change for fathers with high ACEs was very large ($d=1.20$) and statistically significant ($d=1.20$, $p=.05$).

Child outcomes—Children's mean behavior problem scores decreased from 9.25 ($SD=3.55$) at baseline to 8.33 ($SD=2.42$) at post-test, indicating approximately one fewer behavior problem reported. This change was small and statistically insignificant ($d=-.25$, $p=.41$). This was also true for fathers with low ACE scores ($d=-.09$, $p=.83$). However, fathers with high ACEs reported large and statistically significant reductions in child problems ($d=-1.80$, $p=.02$).

Fast-Cycle Learning to Maximize Impact of Subsequent FIND-F Implementation

Analyses of feasibility data showed that FIND-F was attractive to fathers and practical to implement. There was a high completion rate and, on average, fathers completed all six sessions within the intended timeframe. However, it should be noted that this study focused on resident fathers, 60 percent of whom were married, unlike most prior father interventions, which have focused largely on non-resident, unmarried fathers. There was also some indication that the home visiting program model would have to be modified to incorporate evening and weekend sessions in order to successfully implement FIND-F for fathers on a larger scale.

Analyses of pre-post data supported many relations proposed in the TOC. Specifically, FIND-F was related to improvements in fathers' observed parenting skills, which were the primary intervention target. In addition, all fathers showed decreases in parenting stress, and fathers who had experienced high levels of childhood adversity also showed gains in psychological and behavioral involvement and decreases in their child's behavior problems. These findings are being shared across the full Innovation Cluster to identify potential combinations of strategies that together can deliver the largest benefits for identified subgroups of children and families. For example, FIND-F may be combined with other interventions in the Cluster (e.g., those targeting maternal parenting at earlier or later stages of development) to maximize its impacts. As sites increasingly test multiple strategies within their programs (see Figure 1), hypotheses about combined approaches will be tested.

Additionally, if findings from this micro-trial are replicated, FIND-F may be one targeted approach for improving outcomes for fathers who have experienced significant adversity in their own lives and for their children, whereas other strategies being tested in the Cluster may be more beneficial for use with other subgroups of fathers. Given that FIND-F is a one-on-one intervention offered in a home visiting context, identifying targeted populations that will be most likely to benefit from the intervention is critical for scaling.

FIND is also currently being tested with other populations (e.g., mothers and teachers) in a variety of contexts (e.g., child care centers and home-based child care) in order to better understand how, for whom, and under what conditions its active ingredients produce positive outcomes for children. Children's Home Society of Washington has been a partner in many of those efforts, and work is currently underway to significantly increase its capacity to deliver the FIND model. This expansion is partially being driven by the Washington State Department of Early Learning as part of their Quality Rating and Improvement System. Within this context, FIND is being implemented statewide with 240 providers of infant and toddler care, thereby demonstrating the state's commitment to scaling effective strategies developed within the Cluster. Building the capacity of Children's Home Society to serve as a

local FIND hub will also increase the viability of scaling FIND-F within other home visiting programs across the state.

Finally, the original pilot testing team is using the findings of this initial micro-trial to inform the intervention materials, TOC, and evaluation plans for two new studies. The first, a randomized controlled trial with fifty Mexican American fathers, will include a 6-month follow-up to explore whether effects are sustained over time. The second, based on the finding that FIND-F is particularly promising for men who experienced adversity in their own childhood, will involve a micro-trial with fathers involved in the child welfare system. These subsequent studies will build on the initial evaluation plan by incorporating measures of fathers' and children's (when age appropriate) executive function skills to test the underlying mechanisms hypothesized in FIND-F's theory of change.

Discussion

Multiple barriers to innovation stifle the capacity of early childhood policies and programs to achieve significantly better outcomes for children facing adversity. Practical and ethical issues constrain experimentation and funding pressures encourage selective reporting of positive results, thereby limiting what can be learned from approaches that are less effective than expected. In program evaluation, a primary focus on whether an intervention "works" fails to address the more important questions about which children and families benefit most and which benefit least (or not at all) and why. In experimental research, requirements for predetermined study designs and fixed protocols for data collection make it difficult to change course based on early findings, while the often slow pace of the peer-review process and its bias against publishing negative results thwart the rapid-cycle sharing, iteration, and learning from failure that drive breakthrough change in other fields.

In this paper we discuss alternative strategies from the ways in which early childhood programs for disadvantaged children have been designed, implemented, and evaluated for half a century. The essential features of this approach include: (1) formulating explicit theories of change based on enhanced understanding of the effects of adversity on development; (2) designing new interventions based on specific hypotheses generated by these theories; (3) engaging in initial testing of new strategies through micro-trials with small numbers of children, parents, and practitioners in different settings to facilitate rapid modification and fast-cycle sharing based on who appears to be benefitting from specific interventions and who does not. This new approach is illustrated by the creation of a statewide Innovation Cluster within the rapidly growing Frontiers of Innovation learning community and the specific micro-trial testing of the FIND-F video coaching intervention for fathers.

The model described in this paper is based on the assumption that breakthrough outcomes and population level change for children facing adversity will not result from the discovery of a single "evidence-based program" but rather from a portfolio of approaches tailored to the distinctive assets and needs of different children and families. This paper presents the early experiences and lessons learned from a diverse network of practitioners, researchers, service recipients, and policymakers engaged in a disciplined process of fast-cycle design,

testing, evaluation, and sharing of new ideas in real-world contexts. These teams of innovators are supported by a measurement and evaluation infrastructure that facilitates *both* individual learning *and* group-level collaboration. Through this process, knowledge is built both “bottom up,” from the experiences and ideas of those working with children and families in the field, and “top down” through evolving theories of change shaped by scientific insights and empirical data. These strategies bring a much-needed, fresh approach to the evaluation of early childhood interventions that increases our ability to move toward targeted scaling at a faster pace than achieved by conventional research and practice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources: Bezos Family Foundation, Bill and Melinda Gates Foundation, Hemera Foundation, Washington State Department of Early Learning

References

- Anderson S, Roggman LA, Innocenti S, Cook G. Dads' Parenting Interactions with Children: Checklist of Observations Linked to Outcomes (PICCOLO-D). *Infant Mental Health Journal*. 2013; 34:339–351. DOI: 10.1002/imhj.21390
- Bronfenbrenner U. Toward an experimental ecology of human development. *American Psychologist*. 1977; 32:513–531. DOI: 10.1037/0003-006X.32.7.513
- Bruce J, Gunnar MR, Pears KC, Fisher PA. Early adverse care, stress neurobiology, and prevention science: Lessons learned. *Prevention Science*. 2013; 14:247–256. DOI: 10.1007/s11121-012-0354-6 [PubMed: 23420476]
- Cabrera NJ, Shannon JD, Tamis-LeMonda C. Fathers' influence on their children's cognitive and emotional development: From toddlers to pre-k. *Applied Developmental Science*. 2007; 11:208–213. DOI: 10.1080/10888690701762100
- Camilli G, Vargas S, Ryan S, Barnett S. Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*. 2010; 112:579–620.
- Center on the Developing Child at Harvard University. The Foundations of Lifelong Health are Built in Early Childhood. 2010. <http://www.developingchild.harvard.edu>
- Center on the Developing Child at Harvard University. Building the brain's “air traffic control” system: How early experiences shape the development of executive function: Working paper No.11. 2011.
- Chamberlain P, Reid JB. Parent observation and report of child symptoms. *Behavioral Assessment*. 1987; 9:97–109.
- Coburn, CE., Penuel, WR., Geil, KE. Research-practice partnerships: A strategy for leveraging research for educational improvement in school districts. William T. Grant Foundation; New York, NY: 2013.
- Cowan, CP., Cowan, PA. Couple Communication Questionnaire. Berkeley: Institute of Human Development, University of California; 1990.
- Cowan, CP., Cowan, PA. The Pie. In: Touliatos, MSJ, Perlmutter, BF., Straus, MA., editors. *Handbook of family measurement techniques*. Newbury Park, CA: Sage; 1991. p. 278–279.
- Cowan PA, Cowan CP, Knox V. Marriage and fatherhood programs. *The Future of Children*. 2010; 20:205–230. DOI: 10.1353/foc.2010.0000 [PubMed: 20964138]
- Duncan GJ, Magnuson K. Investing in preschool programs. *The Journal of Economic Perspectives*. 2013; 27:109–132. DOI: 10.1257/jep.27.2.109 [PubMed: 25663745]

- Fisher PA, Berry M, Noll LK, Yockelson M, Frenkel T. A model scalable intervention based on developmental translational neuroscience: The FIND video coaching program. *Child Development Perspectives*. in press.
- Fox SE, Levitt P, Nelson CA. How the timing and quality of early experiences influence the development of brain architecture. *Child Development*. 2010; 81:28–40. DOI: 10.1111/j.1467-8624.2009.01380.x [PubMed: 20331653]
- Gorey KM. Early childhood education: A meta-analytic affirmation of the short- and long-term benefits of educational opportunity. *School Psychology Quarterly*. 2001; 16:9–30. DOI: 10.1521/scpq.16.1.9.19163
- Gormley WT, Phillips D, Gayer T. The early years: Preschool programs can boost school readiness. *Science*. 2008; 27:1723–1724. DOI: 10.1126/science.1156019
- Gottfredson DC, Cook D, et al. Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science*. 2015; 16:893–926. DOI: 10.1007/s11121-015-0555-x [PubMed: 25846268]
- Gunnar MR, Fisher PA. Bringing basic research on early experience and stress neurobiology to bear on preventive interventions for neglected and maltreated children. *Development and Psychopathology*. 2006; 18:651–677. DOI: 10.1017/S0954579406060330 [PubMed: 17152395]
- Howe GW, Beach SR, Brody GH. Microtrial methods for translating gene-environment dynamics into preventive interventions. *Prevention Science*. 2010; 11:343–354. DOI: 10.1007/s11121-010-0177-2 [PubMed: 20422290]
- Julion WA, Breitenstein SM, Waddell D. Fatherhood intervention development in collaboration with African American non-resident fathers. *Research in Nursing & Health*. 2012; 35:490–506. DOI: 10.1002/nur.21492 [PubMed: 22685066]
- Loyd BH, Abidin RR. Revision of the Parenting Stress Index. *Journal of Pediatric Psychology*. 1985; 10:169–177. DOI: 10.1093/jpepsy/10.2.169 [PubMed: 4020601]
- MacKinnon DP, Luecken LJ. How and for whom? Mediation and moderation in health psychology. *Health Psychology*. 2008; 27:S99.doi: 10.1037/0278-6133.27.2 [PubMed: 18377161]
- National Scientific Council on the Developing Child. Young children develop in an environment of relationships. 2004. Working Paper No.1
- National Scientific Council on the Developing Child. Excessive stress disrupts the architecture of the developing brain: Working paper No. 3. 2005/2014. Updated edition
- National Scientific Council on the Developing Child. The science of neglect: The persistent absence of responsive care disrupts the developing brain: Working Paper No.12. 2012.
- Patton, MQ. Qualitative research and evaluation methods. Sage Publications; Thousand Oaks, CA: 2002.
- Rothman AJ. Exploring connections between moderators and mediators: Commentary on subgroup analyses in intervention research. *Prevention Science*. 2013; 14:189–192. DOI: 10.1007/s11121-012-0333-y [PubMed: 23007850]
- Shonkoff JP. Leveraging the biology of adversity to address the roots of disparities in health and development. *Proceedings of the National Academy of Sciences*. 2012; 109:17302–17307. DOI: 10.1073/pnas.1121259109
- Shonkoff JP, Fisher PA. Rethinking evidence-based practice and two-generation programs to create the future of early childhood policy. *Development and Psychopathology*. 2013; 25:1635–1653. DOI: 10.1017/S0954579413000813 [PubMed: 24342860]
- Shonkoff JP, Garner AS. Committee on Psychosocial Aspects of Child and Family Health, Committee on Early Childhood, Adoption, and Dependent Care, Section on Developmental and Behavioral Pediatrics. The lifelong effects of early childhood adversity and toxic stress. *Pediatrics*. 2012; 129:e232–e246. DOI: 10.1542/peds.2011-2663 [PubMed: 22201156]
- Shonkoff, JP., Phillips, DA., editors. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, D.C: National Academy Press; 2000.
- Tikunoff WJ, Ward BA. Collaborative research on teaching. *The Elementary School Journal*. 1983; 83:453–468.

- Wagner J. The unavoidable intervention of educational research: A framework for reconsidering researcher-practitioner cooperation. *Educational Researcher*. 1997; 26:13–22. DOI: 10.3102/0013189X026007013
- Yoshikawa, H., Weiland, C., et al. Investing in our future: The evidence base on preschool education. Society for Research in Child Development; 2013.

Figure 1.
Washington State Innovation Cluster: Research Institutions, Intervention Strategies, and
Innovating Sites

Figure 2.
Common Theory of Change Template

Figure 3.
FIND-F Theory of Change

Table 1

Results of Paired Sample T-Tests Measuring Change from Pretest to Posttest

<i>Intervention Target</i>	Measure	Pretest <i>M</i> (<i>SD</i>)	Posttest <i>M</i> (<i>SD</i>)	<i>t</i> (df)	<i>d</i>	<i>p</i>
Parenting Skills	PICCOLO					
Total sample		31.30 (6.92)	34.90 (5.36)	2.85 (9)	.90	.02
Low ACEs		30.75 (8.31)	35.17 (6.42)	2.11 (5)	.99	.06
High ACEs		32.13 (5.20)	34.50 (4.12)	1.38 (3)	.69	.26
Father Outcomes						
Parenting stress	PSI-SF					
Total sample		73.27 (15.39)	63.08 (13.95)	-4.54 (11)	-1.31	.001
Low ACEs		69.89 (19.61)	62.00 (15.72)	-3.56 (6)	-1.35	.01
High ACEs		78.00 (5.15)	64.60 (12.64)	-3.12 (4)	-1.39	.04
Behavioral involvement	WDW					
Total sample		3.69 (.99)	3.89 (1.14)	1.48 (11)	.44	.17
Low ACEs		3.73 (1.10)	3.76 (1.26)	.20 (6)	.07	.85
High ACEs		3.63 (.95)	4.08 (1.06)	2.14 (4)	.96	.09
Psychological involvement	PIE					
Total sample		125.58 (64.50)	143.92 (64.54)	1.72 (11)	.50	.11
Low ACEs		134.00 (44.94)	132.43 (30.57)	-.21 (6)	-.08	.84
High ACEs		113.80 (84.10)	160.00 (97.46)	2.68 (4)	1.20	.05
Child Outcome						
Problem Behaviors	PDR					
Total sample		9.25 (3.55)	8.33 (2.42)	-.85 (11)	-.25	.41
Low ACEs		9.14 (4.63)	8.71 (2.81)	-.23 (6)	-.09	.83
High ACEs		9.40 (1.52)	7.80 (1.92)	-4.00 (4)	-1.80	.02