Volume 31, Number 1 | 2018

Social Policy Report

Toward High-Quality Early Childhood Development Programs and Policies at National Scale: Directions for Research in Global Contexts

Hirokazu Yoshikawa and **Alice J. Wuermli,** New York University
Abbie Raikes, University of Nebraska
Sharon Kim and **Sarah B. Kabay,** New York University

ABSTRACT

With decades of evidence to support early childhood development (ECD) programs and policies, investment in ECD has expanded worldwide. Currently, over 70 nations have national ECD legislation, the majority in the last 20 years. However, with these increased investments comes evidence that the capacity of policy systems to support ECD—across health, education, social protection, and other sectors—is weak, with unfulfilled developmental potential a serious consequence within and across countries. This report aims to develop a research agenda on the systems-level factors—at national, subnational, and local or municipal levels—that may enable or constrain program site-level implementation. Two types of scale—"small to bigger" and "big to better"—are described, as well as the specific challenges of these processes in the field of ECD. Systems factors are reviewed at the three levels, with implications of each for measurement. Finally, methodological challenges and directions are discussed with the aim of informing a research agenda to support national policy progress in early childhood development.

Author Note:
Work on this paper was supported by funding from the NYU Abu Dhabi Research Institute to the Global TIES for Children Center at New York University. We thank J. Lawrence Aber, Harold Alderman, Maria Caridad Araujo, Jere Behrman, Alejandro Ganimian, Liana Gertsch, Aisha Yousafzai and participants at the meeting "Evidence for Scale: Research Methods to Support the Implementation of Quality Education and Early Childhood Development Programs at Large Scale in Global Contexts," held at NYU Abu Dhabi in March 2016. Correspondence concerning this article should be addressed to Hirokazu Yoshikawa, Global TIES for Children Center, New York University, 726 Broadway, 5th Floor, hiro.yoshikawa@nyu.edu.

Social Policy Report

Volume 31, Number 1 | 2018
ISSN 1075-7031

Social Policy Report
is published four times a year
by the Society for Research in
Child Development.

FROM THE EDITOR

This *Social Policy Report* is both international in its perspective and pragmatic in its research recommendations. It is an argument for the development of new research to evaluate Early Childhood Development Programs (ECDs) both in the United States and around the world. The authors pose the question: What sort of research is needed to declare ECDs successful?

Early Childhood Development programs now exist in over 70 countries and have been endorsed as part of the United Nations Sustainable Development Goals recommendations. The question addressed in this report is: What systems level factors can be prioritized for measurement and research in order to facilitate "small to bigger" and "big to better" processes to achieve quality of national scale in ECD services? In short, can we develop the expertise to take ECD programs that have been demonstrated to be effective at a smaller scale to the national level? What sort of research is needed in order to do so?

As the authors point out, what is needed is a new approach to research apart from those that evaluate the success of an ECD at the individual level, or rather, studies the effects of a particular ECD on the children and families who participate in the program. The approach proposed here is to go beyond individual-level data collection and offer an analysis of how to conduct studies that will aid national implementation of ECDs. The authors provide three frameworks, or fields of study, which might be used in conducting research at national or international levels:

1. A dynamic systems theoretical approach, which focuses strongly on the interrelationships among stakeholders involved. Systems theory examines the various factors at multiple levels of analysis that influence the workings of complex systems. Developmentalists are familiar with Bronfenbrenner's model of human development, which utilizes systems theory. Of importance for this discussion are the ways in which various factors at different levels of analysis (individual, community, national, international) interact to influence the success of ECDs and what constitutes success at different levels.
2. The second framework is the implementation science approach for evaluating the effectiveness of an intervention such as an ECD program. Here, researchers evaluate the policies, financing, organizational partnerships, leadership, monitoring, and other factors that might increase the efficiency in implementing programs. This framework focuses attention on incentives for behavior change that would lead to implementation success.
3. Lastly, the authors propose a sociopolitical and cultural theoretical framework, which highlights the political and cultural processes involved in implementation when taking programs to scale nationally. Both traditional policy analysis involving an examination of stakeholders' interests and performance, as well as a more nuanced view of cultural differences (which might affect a country's implementation of an ECD), are considered in this third framework.

This report offers a variety of examples of how to use these frameworks in conducting research to enable the best, highest quality, national-level implementation of ECD programs demonstrated to positively impact children's health and development. Further, this SPR offers more than a guideline to future research; it offers an important contribution to the international literature on early childhood programs and policies.

The conceptualization of what it means to take programs to "scale" offered here (i.e., "small to bigger" and "big to better") are useful rubrics that should engender a more expansive view of program expansions at the national and even international levels. Yoshikawa et al.'s discussion of what data need to be collected and how such data might be used for implementation purposes is clear and compelling. This *SPR* offers practical insights for both policy practice and policy related research.

Toward High-Quality Early Childhood Development Programs and Policies at National Scale: Directions for Research in Global Contexts

Decades of developmental, economic, and program evaluation research have resulted in a growing global consensus regarding the value of investing in early childhood development (ECD) (Black et al., 2017; Shonkoff & Phillips, 2000; Young, 2003). In part as a result of this cumulative evidence base, close to 70 countries currently have national ECD policies, and this number continues to rise (Richter et al., 2017; Vargas-Barón, 2015). The United Nations recently incorporated "quality early childhood development, care and education" for all children prominently in the 2015-2030 Sustainable Development Goals (SDGs). This development expanded the global policy agenda in ECD beyond infant and maternal mortality to include indicators for both universal access to preprimary education and child "health, learning and psychosocial well-being" (United Nations, 2017). This is likely to spur further national and international resource mobilization for ECD programming (ECD Action Network, 2016).

Despite these increases in investments in early childhood development in recent years, ensuring high quality ECD services at scale—making sure that all children have access to quality ECD services targeted to their levels of need—remains a huge challenge. It is estimated conservatively that close to 250 million children, or one in three children worldwide, are at risk of failing to meet their developmental potential due to extreme poverty or stunting (Black et al., 2017; Lu, Black, & Richter, 2016). While there are evidence-based interventions available to support children's development, the challenge this presents is in many ways defined by the issue of scale.

The ability to achieve the SDG agenda as it relates to ECD will depend in large part on the development of strong systems to support effective implementation of policies and practices.

The ability to achieve the SDG agenda as it relates to ECD will depend in large part on the development of strong systems to support effective implementation of policies and practices across international, national, subnational, and municipal levels. A recognition of complex systems is central to progress on national sustainability goals (Cooley & Ved, 2012; Wiek et al., 2012). Though this is true for many of the SDGs, it is a particular concern for ECD given the field's relative lack of systems-level research and the current limited extent of integrated, high-quality services at national scale. For the most part, research on ECD has focused on assessing individual child development or setting-level factors (e.g., in home- or center-based services; the microsystem level in Bronfenbrenner's well-known heuristic; Bronfenbrenner & Morris, 2006). Scale involves more than the simple addition of individuals or sites to particular programs, and therefore research must encompass systems-level factors beyond the individual and microsystem (Myers, 1984; Yoshikawa, Rosman, & Hsueh, 2002). Moving towards national systems and universal, high quality programs will thus require great effort on the part of both ECD researchers and practitioners.

Two Pathways to National Scale: "Small to Bigger" and "Big to Better"

There are two categories of processes leading to accessible, quality ECD services at the national level. "Small to bigger" represents instances when small-scale pilot initiatives "go to scale" and are expanded to cover broader populations. Scaling discussions in rich countries concerning implementation of evidence-based programs, for example, sometimes assume this process, whereby small-scale programs are first evaluated in efficacy trials and gradually moved to widespread adoption and implementation (Coie et al., 1993; Supplee & Metz, 2015). Demonstration projects with high quality at small scale have constituted much of the evidence base on which ECD policy rests (e.g., the Jamaica Roving Caregivers program, which has now developed materials for country-level adaptation and scale; Gertler et al., 2014; Grantham-McGregor, Powell, Walker & Himes, 1991). However, several studies in the global evidence base have shown that when smaller-scale programs go to a larger scale, positive effects are often diminished or lost (for example, Bold, Kimenyi, Mwabu, Ng'ang'a, & Sandefur, 2013; Dodge, 2011). On the other hand, some recent examples show success when the process of scaling involves multiple stages and testing in a variety of local policy and implementation contexts—that is, when across iterations the core of a program's active ingredients are maintained in adaptations to larger systems-level factors, such as workforce capacity, local policy, supervision and governance (Banerjee et al., 2016). "Small to bigger" expansion may be more successful if planning for scale occurs starting during the phase of small-scale implementation (Dodge, 2011; Supplee & Metz, 2015). However, this generally implies that the program being tested has already been shown to be effective in another context. This may involve, for example, both research on and engagement of a wider set of systems-level characteristics and actors than one would if small-scale implementation were not expanded.

"Big to better" refers to systems that are already at scale but delivering services that require some improvement. For example, health care systems providing antenatal care, immunizations or treatment, or social protection systems which provide income and material support to certain populations, are already at national scale in many countries. This second approach addresses a significant challenge to scaling in that the quality of services at scale in low- and middle-income countries is often quite low, and in some cases worse than the counterfactual of no services (for example, in the case of public center-based care in some countries in Latin America; Rosero & Oesterbeek, 2011). Improvement may also be required for accurate targeting of services (e.g., screening and data systems to identify the disadvantaged for receipt of cash transfers; screening caregivers for depression to receive additional mental health intervention). Effective ECD program provision in at-scale systems must also recognize diverse needs across heterogeneous populations (Schindler, Fisher, & Shonkoff, 2017). In this case, when diversity of needs has not been considered from the start, one will need to go back and identify the core of a program and which active ingredients need to be strengthened and what the contextual factors are that undermine or support quality service delivery.

The Problem, Aim, and Definitions

A central problem now faces the field of early childhood: How to expand or improve effective practices across populations so that all children who need them receive access

to quality early childhood experiences. At the heart of this problem is how to 1) build systems to support large-scale service provisions in places where such systems do not exist; and 2) how to change behavior across existing, but low-functioning, systems to improve quality. These two problems require new approaches to scaling, as both types of problems have been shown to be stubborn and complex in existing studies.

This paper aims to identify critical factors in the study of quality ECD services and systems at national scale. Rather than cataloguing case examples of successful attempts at scale (for recent examples see Bernard van Leer Foundation, 2011; Richter et al., 2017; Robinson, Winthrop, & McGivney, 2016), our objective is to develop a research agenda for system-level factors and processes to inform the two major kinds of scale. Therefore, our central research question is: What systems-level factors can be prioritized for measurement and research in order to facilitate "small to bigger" and "big to better" processes to achieve quality at national scale in ECD services? As more and more countries engage in national ECD policies and attendant expansion of ECD programs, we believe a research agenda to assess these forms of scale will serve as an important supplement to the traditional impact evaluation methods that generally define ECD program and policy research.

In this paper, we define "scale" to include processes at three population and governance levels: national, subnational, and local or municipal levels of jurisdiction, including all children and families living in a jurisdiction defined by the national governance structure. We thus prioritize policy-oriented levels of scale, given our initial frame of the national ECD policies and associated programs that may realize progress on SDG Target 4.2 and related targets and goals. We emphasize all children, similarly, due to the universal nature of the SDGs, but also discuss issues related to the targeting of services. Targeting is a critical concept for the field of ECD, especially in relation to conceptualizations of scale in the context of low-resource contexts. We exclude factors that are primarily at levels below the municipality—for example, the single-program-site level (and therefore the microsystem level; Bronfenbrenner & Morris, 2006). Although efforts to achieve scale aim to support program-level features, we do not review all the program-level features that are considered important in ECD research (these are covered in several recent reviews, including Britto, Yoshikawa, & Boller, 2011; Pianta, Downer, & Hamre, 2016; Sun, Rao, & Pearson, 2015). Finally, whereas some related recent work has looked at scaling evidence-based programs with a focus on the United States (e.g., Supplee & Metz, 2015), we take an explicitly global focus, with examples from both high-income (HIC) and low- and middle-income countries (LAMIC), in keeping with the global nature of the SDGs.

We begin by introducing three strands of research and resulting paradigms that we believe are relevant to the understanding of "small to bigger" and "big to better" types of scale: dynamic systems theory, implementation science, and sociopolitical theory. These three strands of research exhibit substantial commonalities, but have been pursued from different disciplinary bases, and target different levels of the system. We argue that research needs to draw on all three in order to account for a complete systems approach to studying quality and scale in ECD. We then apply these lenses and frameworks to the specific challenges of the ECD field. This application will provide the basis of a taxonomy of prioritized systems-level factors aimed at building an evidence base on "small to bigger" and "big to better" processes in ECD policy.

Drawing on Existing Analytic Approaches in the Study of Scale and Quality in Social-Service Sectors

Quality and impact at scale are central to social-policy sectors. Advances in conceptualizing and measuring large-scale service and policy implementation are occurring simultaneously in several social-science and applied disciplines (across, for example, research in poverty, public health, nutrition, mental health, and education). Although many fields have noted the difficulty in relying on purely additive or linear models to explain the differences between site-level and at-scale implementation, no single consensus framework for scale in social-sector programming currently exists (Fixsen, Blasé, Naoom, & Wallace, 2009). Research on this front is most robust in the healthcare sector, which has had a longstanding focus on issues of scale and quality in systems. It is also prominent in the organizational systems literature (Aarons, Hurlburt, & Horwitz, 2011; Cooley & Ved, 2012; Shiffman & Smith, 2007).

What we know from some other sectors, such as education, is more limited—for example, a review of studies of international development with a focus on scaling found that only 16 of 158 include a focus on education (Pidufala, 2008). Some recent initiatives are relevant to scale in child development and education policies, such as the Research on Improving Systems of Education (RISE) Programme—a large, global collaboration of organizations and researchers that seeks to answer questions around how to improve education systems to deliver quality learning at scale in low- and middle-income countries (e.g., Pritchett, 2015). A recent report from the Brookings Institution (Robinson et al., 2016) also identifies core ingredients that contribute to scaling quality education using a comparative case study approach. Here we highlight common characteristics from three of the most prominent fields of inquiry: dynamic systems theory, implementation science, and sociopolitical theories.

Dynamic Systems Theory

Dynamic systems theory (also referred to as complexity theory, systems theory, and other labels) is one influential approach to the study of complex environmental and social processes and systems, both formal and informal. Dating back to the mid-20th century (e.g., Von Bertalanffy, 1968; Waddington, 1957), dynamic systems theory has incorporated several principles of the functioning of complex systems. These include: 1) how factors at multiple levels function jointly, 2) how factors at multiple levels can influence one another, 3) how factors at multiple levels can have interdependencies, and 4) how factors at multiple levels can operate and influence one another in non-linear ways (de Savigny & Adam, 2009, Kroelinger et al., 2014). Bronfenbrenner's model of human development may be considered one of the earliest applications of systems theory to child development (Bronfenbrenner, 1979). Dynamic systems theory has since gained prominence in the developmental sciences (Ford & Lerner, 1992; Thelen & Smith, 1994; Masten, 2007), though only in a few instances has it been considered or applied to large-scale human developmental programs and policy implementations (Lundberg & Wuermli, 2012; Yoshikawa & Hsueh, 2001).

A central tenet of this theory is that a system is composed of multiple interconnected parts that will naturally work to reach a new equilibrium when change is introduced, through and within the connections between various parts of the system. Therefore,

a system cannot be understood by simply breaking it down into its component parts. Understanding a national home visiting program, for example, cannot be accomplished by simply applying measurements of parent-child interactions or even home visitor-parent or trainer-home visitor interactions to more and more instances. According to dynamic systems theory, aggregated individual-level behavior cannot by itself explain or predict systems-level behavior, as systems have their own properties and characteristics at higher levels. Rather, systems theory hypothesizes that behavior results from multiple levels of interaction between individuals and systems, and maintains that lasting behavior change is instigated through frequent and simultaneous reinforcement of small changes at multiple levels of the system that in turn can lead to sequences of either incremental or discontinuous change. This perspective is important for both the "small to big" and "big to better" types, in that transitioning a small-scale intervention to a larger scale must adapt the intervention in order to address the different demands of larger scale systems.

Dynamic systems theory also suggests that there are multiple entry points through which a system can be changed and improved (Wuermli et al., 2012). Therefore flexibility or adaptability of approaches for scaling depend on the geographic, political, economic, and sociocultural context. For example, one might provide specificity in national quality standards, enhance local input into definitions of quality, or strengthen subnational institutions for monitoring of quality. Decisions about level at which to intervene are likely to be influenced by availability of resources, time constraints, and other political and social factors specific to a particular country's context.

As systems thinking is applied in social sciences, sometimes the theory has been used to describe the necessity of building systems to support scaling. Rather than approaching problems with the mindset of scaling an intervention using an additive or linear approach (e.g., simply adding sites), systems thinking works to uncover the characteristics and relationships that undergird systems. This approach, as it has been applied to health systems, has provided insights on how to expand service sites through engagement of higher-level complex systems and real-world settings of services and, ultimately, how to improve outcomes for children and families (de Savigny & Adam, 2009). For example, reciprocal feedback loops in which information is shared not only from higher policy levels to lower ones, but also from lower to higher, are encouraged in both continuous quality improvement and recent scorecard methods to reducing errors in health care provision (Arbour et al., 2015; Berwick, 2003; Lingard et al., 2008). The systems-theoretical principle of "tipping points," in which systems demonstrate discontinuous change once a certain parameter is reached in a more gradual process, has informed efforts to estimate proportions of populations to target for behavior change to spread across the entire population (Trochim et al., 2006).

A recent application of systems theory for the education sector has been proposed by Pritchett (2015). He argues that the Millennium Development Goal and Education for All global initiatives produced systems designed to increase access rather than learning. In order to improve actual student learning, he argues that the inter-relationships of major components of educational systems—from national policy frameworks to subnational and teacher workforces, to the client population of students—must create data-based feedback loops that are based on learning rather than enrollment.

This critique is also highly relevant to the field of ECD and the subfield of early childhood education (ECE). There have been notable increases in access to pre-primary education, which may be reinforced by the emphasis on pre-primary enrollment as a key global indicator of ECE. Singular efforts to increase enrollment in preschool run the risk of replicating many of the problems facing the primary education sector. Rather, a set of policy levers—teacher incentives and training, population demand, quality monitoring, and data-based accountability, for example—each with their particular characteristics and constraints, may each be relevant, but act in interdependent fashion to determine the ultimate effectiveness of ECD policies in enhancing child development at scale (Vegas & Ganimian, 2011). Rather than repeat the primary education sector's focus on enrollment during the MDG era, ECE must consider a range of systems factors that determine child learning and development aside from access.

Implementation Science

Implementation science is an emerging field of study with roots in global public health. In the United States it has found its way into implementation frameworks in ECD (Metz, Naoom, Halle & Bartley, 2015). Most of implementation science focuses on program-level dimensions of fidelity such as dosage, quality of services provided, and participant response to services as contributing factors to effectiveness (Domitrovich et al., 2008; Hulleman & Cordray, 2009). The portion of this research area most relevant for our purposes concerns influences that support successful, high-quality implementation at large scale. Some models have focused on measuring and adjusting evidence on targets and mechanisms of programs to a wider set of local contexts or considering the workforce supports for local program quality that can promote expansion with fidelity of programs previously evaluated at small scale (Aarons, Hurlburt, & Horwitz, 2011; Domitrovich et al., 2008; Fixsen et al., 2009; Supplee & Metz, 2015). A key focus of implementation science is behavior change, or how to influence systems to provide resources and supports for changes in behaviors that will lead to better practices over time. Such a model has clear implications for improving quality in large-scale systems.

Unlike more general systems theory, implementation science often focuses on preexisting models (e.g., evidence-based practices or evaluated curricula) that are then expanded (Kroelinger et al., 2014). A sequence of pre-adoption, adoption, implementation, and sustaining of programs refers to the process of embedding a preexisting program model into large-scale systems. Nadeem and colleagues (2014) provide a summary of critical factors from across several implementation models, including systems factors external to the program site level such as policies, financing, organizational partnerships, and leadership that can foster this sequence of adoption and quality implementation. Aarons et al. (2011) apply a dynamic aspect in which larger-scale systems interact over time in a mutually reinforcing manner with lower-level program site-level implementation. They also state that the fit between an evidence-based program and a larger system is an important consideration.

When implementation science is applied to the "big to better" approach, efforts to increase efficiency or develop better monitoring systems can lead to improvements in the implementation of services that already exist at large scale. Studies on improving the quality of care provided have been conducted on a new national policy regarding

post-natal care in Ghana (Twum-Danso et al., 2014). Global health scholars have called for implementation science when testing solutions for quality improvements reaching from macro-factors (e.g., provider training, health insurance, accountability mechanisms) to meso-level factors (e.g., district management and supervision) to micro-factors (e.g., clinic level factors like checklists and supervision; Kruk, Larson & Twum-Danso, 2016). The Institute for Healthcare Improvement's model of continuous quality improvement specifies both global systems-level indicators of quality of health-care systems and lower level indicators defined and measured by networks of stakeholders that can in turn predict quality practices at the clinic level, like quality checklist implementation (Institute for Healthcare Improvement, 2007). Thus, work in public health has linked program-level implementation indicators to systems factors.

Sociopolitical and Cultural Theory

Political and cultural factors are not often explicitly included in systems frameworks. Political theories have traditionally focused somewhat more on the policy creation process rather than the intricacies of policy implementation (see e.g., Skocpol, 1995). However, some have focused on implementation processes in addition, including in the ECD field (Shiffman & Smith, 2007). Politically-oriented theories often create a clearer distinction between national, subnational, and local processes and governance. They also consider the coalitions and their interactions that influence policy across civil society, labor, advocacy organizations, and NGOs, not only governments. In the health literature, Shiffman and Smith (2007) have outlined how political priorities can heavily influence the ability of an evidence-based health intervention to achieve widespread adoption and scale, both from the perspective of a broader enabling environment and from a program-level perspective. In their view, political processes occur not just at the national level, but also at the grassroots level.

Along the lines of these theories, political scientists have long recognized the importance of networks in policy implementation. Policy networks consist of various stakeholders, including political leaders, local governments, businesses, civil society organizations, and communities. Policy network analysis (Rhodes, 1997) involves mapping stakeholders and stakeholder relationships, including the relative power attributed to each player, and has been used to analyze phenomena at district and regional levels. For example, Coburn and Russell (2008) examined the variation in impact of district-level educational policy on students through variation in the functioning and content of teacher social network interactions. Guthrie (2001) examined how leadership networks at the city, industry, and within-city municipal levels in Shanghai carried out the massive economic reforms in the 1980s and 1990s transition from a state- to a market-based economy.

Analysis of political processes can be applied to both "small to bigger" and "big to better," across political/policy networks to institutional and coalition-based analyses informed by political science. These perspectives can help us further understand the social forces that can lead to or constrain the scaling of ECD systems, as well as how to generate and maintain the political will to expand and improve ECD programs and policies.

Colombia created a *modalidad propia*, a process through which indigenous and rural groups could define their own culturally anchored solutions to meeting national program and quality standards in ECD programming.

A strong critique of traditional models of scaling interventions comes from cultural perspectives on human development and intervention science. For example, Brouwers (2017) applies the lens of cross-cultural psychology to critique general models of intervention that are assumed to work similarly across cultural and linguistic groups. Approaches to intervention development that build on local cultural practices rather than the solely Western models of child development that predominate in developmental science have been put forward by Cole (2005) and Serpell (2011). These models have not often incorporated policy perspectives or challenges at national scale. Different countries have taken different directions in balancing universally conceptualized systems characteristics (for example, quality standards in ECD programs) and the autonomy of specific cultural groups and institutions in a society to act within the national policy implementation framework—for example, autonomy to determine the modality and content of such programs. In one well-known example, the New Zealand national preschool curriculum was reconceptualized and centered on Maori cultural values and principles through an inclusive stakeholder process (Carr & May, 1993). In a more recent example, Colombia created a *modalidad propia* (own modality, alongside family and institutional modalities of services), a process through which indigenous and rural groups could define their own culturally anchored solutions to meeting national program and quality standards in ECD programming (Motta & Yoshikawa, in press). This model allows for local definitions of quality alongside national definitions, as called for by Dahlberg, Moss, and Pence (1999), Pence and Marfo (2008), and others.

In sum, conceptual frameworks for large-scale service provision and implementation identify processes that go beyond the individual and the microsystem to larger systems levels. Each of the lenses of systems, implementation, and political and cultural forces is more "macro" in nature than the bulk of research on ECD interventions. Each of these frameworks has components that may be useful for the study of scale in ECD programs and policies (for example, as we have seen, some are more suited to the "small to bigger" form of scale, while others are more suited to the study of the "big to better" type). The ECD sector also has unique characteristics that may inform a research agenda on scale that is specific to the field. We turn to these next.

The Particular Challenge of ECD: Developmentally Informed Sequences of Multisectoral Services

The field of ECD raises particular challenges of implementation for programs and policies at scale. To achieve improvements in ECD at the national level for all children, resources must be coordinated across several sectors with an emphasis on quality and equity, not just access (Black et al., 2017; Britto et al., 2017). Both sector-specific provision and coordination of services are required. Moreover, young children's developmental needs change drastically over the first eight years of life, requiring age-appropriate services.

There are several elements of early childhood development that should be taken into account in a research agenda on scale in the field. First, the task of integrated and coordinated ECD programs and policies involves a high level of multisectoral coordination. Unlike single-sector goals such as universal primary education or universal health care coverage, the goal of maximizing developmental potential across the period of prenatal to age eight (an age range consensus in the ECD field that we adopt in this paper) involves coordination of policies and programs across health, nutrition, education, social protection, and child protection sectors (Richter et al., 2017). As one example, the seemingly straightforward task of improving the quality of responsive adult-child interaction (whether the adults are caregivers, parents, or teachers) is difficult as service providers' contact with parents can cut across all of these sectors. Despite strong evidence that parental responsiveness and stimulation can improve learning and early socio-emotional development, such a focus is rarely incorporated into health, nutrition, social protection, and other services in the first 1,000 days (Britto et al., 2017; Gertler et al., 2014; Yousafzai & Aboud, 2014). And each of these sets of services may serve different and only slightly overlapping populations, as each is rarely at 100% national coverage of its intended population. Thus, to achieve greater coverage of a focus on parenting stimulation and responsiveness, coordination across these sectors is critical.

Second, while also relevant to other services, scaling ECD programs requires an explicit commitment to quality across different government systems, emphasizing the delivery of quality care, health, nutrition, water, sanitation, and education rather than expanding access alone. Despite agreement that quality needs to be prioritized, few systems are truly aligned to promote quality; most are primarily oriented towards access (Pritchett, 2015). Extant data suggests that both the comprehensiveness and quality of services at scale are problems. At the system level, comprehensive service delivery can be challenging in contexts without strong systems infrastructure and human resources. Systems infrastructure refers to the systems to support service delivery (i.e., effective and efficient management and information systems, physical infrastructure, and institutional support structures for human and material resources). For example, social protection systems require accurate management and information systems to assess and target transfers or other resources to their intended populations. Human resources refer directly to the quality of the (para) professional workforce (e.g., skilled health workers or caregivers). For example, the quality of pre-primary education as measured through direct observation is quite low, particularly the levels of teacher-student interactions that promote elaborated conversation, child questions, and guided play (Araujo et al., 2016; Berlinski & Schady, 2015).

Service provision may be public, private, or civil society-based, and funding may come from public, private, or civil society sources. In the early childhood field, for example, the private sector is growing in many country contexts (Woodhead & Streuli, 2013). While alternatives to public provision of services can alleviate supply-side constraints, they introduce another layer of questions regarding regulation, governance, and accountability. Research on scale in ECD needs to take into account the regulatory, oversight, and accountability forces that affect the quality of services differently across diverse providers.

Population-level demand for specific services is particularly low and variable for ECD programs and policies. In many low- and middle-income countries in particular, ECD services have not been at scale long enough that the majority of parents experienced them as children. This is in contrast to some health services or primary education which have achieved near-universal demand. Therefore, demand can be low for even high-quality services in ECD. For example, home visits to support parenting and nurturing care have not been the norm for generations in any country. It may be no wonder then that in most home visiting programs it is rare for targeted populations to take up more than half of the offered number of visits (Doyle, Fitzpatrick, Lovett, & Rawdon, 2015).

A fourth challenge of ECD policies at scale is the need for developmentally-informed sequences of services and associated programs. Few services are continuous in setting or sector across the entire prenatal to age eight period, for example. Services that have attempted to do this, such as center-based care from birth to school entry, as was provided in the landmark Abecedarian demonstration project, have not been scaled up to the national level (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001). Critical health services directed towards child survival, for example, across pregnancy, the first 1,000 days, and the subsequent preprimary years are quite different in nature. Although they can be provided by the same care sector, the roles of more highly trained versus less highly trained workers (nurses and doctors versus community health workers) differs across different services (Richter et al., 2017). Similarly, care and education services vary quite drastically in recommended group sizes, ratios of adults to children, and required caregiving and instructional skills between the first year of life and the preprimary year. To give one example of the consequences for a systems-level construct, there cannot be a single set of quality standards covering all ECD programing between the prenatal and early schooling periods. And social protection policies can differ by developmental period (e.g., paid maternity leave versus more general transfers). The rapid and massive changes that occur in the early years of development ultimately require a particularly complex sequence of recommended programs and services that must differentiate between young children of different ages, with consideration of transitions among them (Kagan & Kauerz, 2012).

In summary, to ensure that early childhood investments at the national level are well-spent, challenges in the field include assessing characteristics of coordinated systems to support required specialized human resources; encouraging comprehensive, high-quality service provision; ensuring that all children and families can and want to access these services; and providing for the very different kinds of programs and services that support each of the earliest periods of human development.

Measurement Approaches for National, Subnational, and Municipal/Local Factors

As mentioned previously, a central part of systems-level analysis is the consideration of a system's multiple levels (Pritchett, 2015). Across all three research perspectives discussed previously (systems, implementation, and political/cultural), there is acknowledgment of the importance of looking at the entirety of a system, which in turn necessitates articulation of a measurement framework. We organize the following framework by national, subnational, and municipal levels, though we acknowledge that some issues are relevant to multiple levels. We summarize these factors in Table 1.

These factors should be considered a heuristic that can serve as the basis for more systematic testing and analyses of processes toward scale in ECD.

Table 1. **Dimensions of Systems-Level Factors: National, Subnational, and Municipal/Local Levels**

NATIONAL LEVEL
National ECD legislation and strategic or action plan
• Existence of national ECD legislation and its multisectorality • Existence of national strategic plan, including budget and quality standards for different service sectors; guidelines for workforce, finance, subnational leadership and governance structure; and processes of review • Participatory level of policy, strategic plan development, and monitoring [proportion of relevant national stakeholder groups consulted and frequency], including involvement of groups across gender, ethnicity, race, indigenous status, language, refugee or migrant status, and other indicators of national diversity • Sustainability of national legislation or budgeting across administrations (e.g., budget permanence) ECD sector
Coordination across sectors
• National coordinating body or lead ministry; regularity of convening for planning • Cross-sector multi-year planning • Cross-sector multi-year budgeting • Accuracy of national data systems to track budgeting, access, quality
Political mobilization
• Extent of networks of political leaders, civil society leaders, and other important stakeholders on behalf of ECD programs and policies
SUBNATIONAL LEVEL
Vertical links and actions based on information from the two other levels
• Accuracy of information passed "up" and passed "down" • Adequacy of resources passed "up" and passed "down," relative to levels necessary to implement programs as intended in policy • Capacity to adjust subnational implementation based on information and resources from higher and lower levels
Characteristics of the subnational workforce
• Definition of competencies and skills specific to the subnational workforce • Training and supervision specific to these competencies
MUNICIPAL OR LOCAL LEVEL
Training and monitoring system characteristics
• Pre-service training availability and alignment with provider skills that predict child development • In-service training availability and alignment with provider skills that predict child development • Fit of expected duties of workforce with other responsibilities, income generation, and daily routines • Monitoring of local variation in resources and compensation adequacy
Municipal-level governance and budgeting
• Municipal-level governance – mayoral or municipal political support, representation, and inclusion • Municipal-level budgeting – relation to actual spending, training, and support for local budgeting roles and skills

Outlining these different levels and the factors assigned to them is meant to illustrate that better understanding system levels will significantly contribute to our understanding of what will define and enable high quality implementation of ECD programs and policies at national scale. The definition and study of levels and their interaction in ECD systems, with a few exceptions, is notably absent from current research on ECD. We use this discussion to make practical suggestions about how "small to bigger" and "big to better" types of scaling can be envisioned, what data and research are needed, and also how the different approaches and methods outlined earlier regarding scale might best apply.

National-Level Factors and Their Measurement

National-level systems factors that could constrain or enable implementation in ECD include the existence of national ECD legislation and a strategic plan linked to budget appropriations, coordination at the national level between sectors, and political mobilization. One global effort exists to measure factors in these categories. The World Bank's Systems Approach for Better Education Results (SABER) framework for early childhood development policies identifies three overarching and interlinked policy domains: 1) establishing an enabling environment, 2) implementing widely, and 3) monitoring and assuring quality. Each has constructs measured within it which are conceptualized as policy levers (Denboba, Hasan, & Wodon, 2015; World Bank, 2016).

In the SABER measurement framework, establishing an enabling environment includes the existence of a legal framework, the extent of intersectoral coordination, and sustained finance mechanisms. Implementing widely includes the three levers of scope of programs, coverage of programs, and equity in their coverage (e.g., inequality in access). Finally, monitoring and assuring quality includes the levers of data availability across national and subnational to individual levels, quality standards and their existence across sectors, and adequacy of a monitoring system for compliance with standards. Each lever is rated on a four-point scale (latent, emerging, established, or advanced). Each can be measured at the national, subnational, and municipal levels. The recent analysis of Indonesia's ECD policies by Dendoba et al. (2015) presents one of the first analyses to supplement these national indicators with their counterparts at provincial and district levels using the SABER framework and data tools. In the following sections we make particular note of dimensions that are not covered by the SABER system.

National legislation and strategic plans. Beyond the very basics of governance, national ECD legislation (whether sectoral or multi-sectoral) and a strategic plan linking the legislation to budget appropriations over a defined period are typical milestones in scaling ECD programs. ECD legislation at the national level (whether a single law or number of laws) is a signal of political will at the time of passage. However, for political will to translate into quality programs sustained over time, many other systems aspects must be in place. The most basic is a strategic plan that translates a national policy into action, with accompanying appropriation of funds. The number of comprehensive national ECD legislation and action plans has been tracked in recent years by Vargas-Baron. She has observed, for example, that only 28 of 61 national ECD laws had accompanying active action plans (i.e., were being implemented; Vargas-Baron, 2015). A strategic plan should incorporate a plan for cross-sectoral (cross-ministerial) collab-

oration and coordination; the establishment of quality standards for services; guidelines for workforce characteristics, competencies and compensation; provision for data systems and regular collection; guidance for subnational leadership; a time frame and periodicity for review and re-authorization; and the primary finance mechanisms for each sector of ECD services (Vargas-Baron, 2005).

The question of approach to scale, "small to bigger" or "big to better," is highly relevant to the discussion of legislation and strategic plans. For countries that are introducing national ECD legislation for the first time, the work might be considered "small to bigger," for example, expanding a program from one or two regions to the national level. Countries that have preexisting legislation and national programs might be more concerned with strategic plans and translating their legislation into action—more of a "big to better" approach. It is important that legislation and planning work strategically to approach the question of scale and are based on accurate and periodic assessment of current conditions.

Several requirements have been hypothesized or proposed for national ECD legislation and accompanying action plans. One is the existence of participatory processes in policy development and implementation. Vargas-Baron (2005, 2015) has written extensively about participatory processes in establishing and implementing ECD policy, the goal of which is to maximize the chances that the policy will be realized in budgeting and other implementation processes. Inclusiveness of stakeholders across national, subnational, and municipal governments; NGOs; civil society members; and representatives of marginalized groups or of diverse ethnicities, languages, and political parties is a principle that applies across all phases of policy development. Phases can include initial national and subnational situation analysis, gathering of data on the status of young children and their families, convening stakeholder groups to determine elements of the national policy, drafting the policy, estimating its costs and major budget elements, obtaining wider input, revising the policy, preparing for its passage, and once passed, developing and monitoring the national action plan for implementation (Vargas-Barón, 2005).

Second, the sustainability of the policy is critical across dimensions of political, financial, and social sustainability. Does an ECD policy survive changes in political parties and contexts? Is financing similarly sustained and adequate to meet objectives from year to year and across political administrations? The challenge of ECD sectors is that they are not typically enshrined in rights-based or other universal entitlements. For example, to ensure universal access to preprimary education, it may be most sustainable for the basic primary education legislation of a country to incorporate access at the preprimary level. Given the pressures to achieve universal secondary education, preprimary may fall to the wayside as basic education does not always include the preprimary level in the form of entitlement-based laws and budgeting. This type of research aligns with the political approach to the study of systems, investigating what types of relationships and contexts enable effective and long-lasting policy-making.

And finally, addressing the role of the private sector and the state's responsibility for monitoring and regulation is often a key part of national legislation. Given limited budgets and the current landscape of ECD services, private sector provision plays a large role in ECD in many countries. The private sector may act as a spur to public

provision by responding to demand more closely and developing innovations from a community-based perspective, while at the same time presenting a challenge to implementation of national policies in areas like quality monitoring.

Coordination across sectors and actors. Coordination at the national level between sectors and service providers involves regular meetings of leadership (minimally, finance, health, education, social protection and child protection, but also transportation, urban and rural development, national statistics and data gathering, and private sector service providers). As has been observed by others (Britto et al., 2014), the coordinating or leadership body in ECD policy differs quite widely across countries. In some cases, the planning agency or ministry coordinates the policy (as in Chile Crece Contigo; Valenzuela, Delpiano, & Cordero Vega, 2011); in others, line ministries coordinate multi-sectoral efforts (e.g., the health or education or social protection ministries); in others a national coordinating body is set up that resides outside the ministries (e.g., in the President or prime minister's office). No single approach appears more successful than others—the power to convene regularly and the ability to influence line ministry policy planning and implementation are perhaps more important than the location of the coordinating body. The existence of such a coordinating body is captured in the SABER system; some researchers also have recorded their locus within national governance systems (Vargas-Baron, 2015). Organizational and related systems theories will be helpful for this analysis, understanding the strengths and weaknesses of different approaches, such as a lead ministry versus an independent coordinating body. Investigating various incentive schemes to encourage coordination and strategies for training on effective coordination will also be constructive.

Political mobilization. Political processes are powerful influences on policy development at the national level, mobilizing the funding, ensuring quality implementation, and ensuring sustainability of the initiative; however, these are not easily recorded and used as data (Shiffman & Smith, 2007). Effective ECD policy action has capitalized on windows of political opportunities: when stakeholders align in support of the issue, moments of consensus for support of particular policy initiatives, or instances when particularly powerful actors lead proposals. Coalition-building across political parties and stakeholders is vital to ensure success of ECD legislation but varies across political systems. The building of political will ultimately aims to build coalitions from diverse and competing interests through social mobilization and organizing, civil-society action, media campaigns, and negotiation at leadership and other levels (Skocpol, 1995). These features are perhaps most difficult to capture quantitatively; case study accounts sometimes incorporate attention to the political mobilization process that led to ECD legislation. Market, cultural, and political approaches to systems research, including policy network analysis (Rhodes, 1997), may be important to understanding these processes. For example, investigating how campaign promises and elections can open up or close policy windows at a national level could be insightful.

Approaches to measurement: Accurate national data systems for spending, access, quality, and child outcomes. National-level data collection on these factors is conducted typically using a key informants approach, as the World Bank SABER

system (Denboba et al., 2015; World Bank, 2016). That approach could be supplemented with some of the data elements described here. This requires several considerations in sampling, including identifying the appropriate administrative level of informants within ministries; identifying the correct periodicity, given rapid turnover in ministry and other national leadership; and addressing challenges in identifying ECD-specific spending within larger expenditure lines.

The actual links between national planning and budgeting processes and subnational and local spending, as well as actual coverage of the national population by ECD services, can only be tracked with accurate reporting mechanisms. Expenditure data collection (if carried out at all) is typically done by ministries' finance and planning departments in coordination with the ministry of finance or treasury and generally does not include private providers. It is usually quite difficult for outside researchers to gain access to these reporting numbers; however, some countries have legislation for "open budgeting" policies that report expenditures in a real-time (or close to real-time) basis. The World Bank's BOOST initiative and associated Open Budgets Portal assists governments to provide the public with national spending data, disaggregated by region and by categories of expenditures. Even with open budgeting, accurate data specific to ECD-related spending can often be difficult to determine since budget lines may include resources for both early childhood and other age ranges or populations (Van Ravens & Aggio, 2008).

Similarly, coverage data must be accurate and supported by data systems that are consistent across regions, provinces, and municipalities and reported at sufficient periodicity to inform ongoing decision-making. Often private providers are not included. Furthermore, the disaggregation of coverage data by dimensions that provide data on equity or inequity—across gender, ability status, ethnicity, language, and rural/urban status—is not consistent, with some dimensions much more consistently reported (e.g., gender; urban/rural) than others (e.g., ability status, which is often not tracked accurately during the early childhood years due to a lack of at-scale screening and diagnostic systems).

Although many countries monitor basic structural features of program quality (facility safety and hygiene, staff qualifications, and attendance) very few monitor aspects of process quality (e.g., quality of interactions in parenting-focused services or in early education).

National data systems should ideally address not only coverage but quality and actual child development and learning (quality is called for in Target 4.2 of the SDGs and a multi-domain indicator of child physical, cognitive, and socio-emotional development in Indicator 4.2.1). For this, consistency and validity of measures across diverse parts of the country are critical. Although many countries monitor basic structural features of program quality (facility safety and hygiene, staff qualifications, and attendance) very few monitor aspects of process quality (e.g., quality of interactions in parenting-focused services or in early education; Berlinski & Schady, 2015).

Recording and tracking the accuracy and quality of the data itself in these data systems is a nascent enterprise in ECD. With the inclusion of indicators of child development

The development of sound, culturally anchored measures and data collection systems is a necessary foundation from which to build scaled systems of ECD.

and learning in the SDGs, global efforts to define technical standards for national data are underway (e.g., through UNESCO's Global Alliance to Monitor Learning; UNESCO Institute for Statistics, 2016). The development of sound, culturally anchored measures and data collection systems is a necessary foundation from which to build scaled systems of ECD. Effective feedback loops and the ability to identify and respond to issues of concern have been identified by systems theory as critical characteristics of high functioning systems (Pritchett, 2015; Thelen & Smith, 1994). This cannot be accomplished without timely and accurate feedback from programs and services, as well as regular population or census-based data.

Subnational-Level Factors and Their Measurement

The subnational level differs across countries not only in the labeling of the units (provinces, states, departments, regions, etc.) but in their functions. In relatively decentralized political systems, subnational units can have comparable or greater influence on large-scale ECD program implementation than the central government (e.g., in Brazil or the United States). In some cases, the national versus subnational roles in planning, budgeting, and governance of ECD services varies by sector. For example, in the United States, infant and toddler nutrition services are implemented through a federal program, while public early childhood education is provided through a mix of programs funded by the federal government and state and municipal governments. In practice, this results in variation in both access and quality supports, as well as monitoring systems, in publicly funded early education in the United States (Chaudry, Morrissey, Weiland, & Yoshikawa, 2017).

The subnational systems of ECD implementation have been the least studied, relative to national processes of policy development and implementation and local systems. One six-country study found that the capacity for cross-sectoral coordination, for example, was weakest at the subnational level, relative to the other two levels (Britto et al., 2014). This was in part because the functions of subnational leadership were often (in their view) not clearly defined and did not include the traditional functions of planning and budget oversight. Rather, the subnational level was often perceived as a pass-through with the main functions of government officials at that level being compliance with national standards and monitoring of lower-level systems. Because monitoring systems are often not linked closely to child outcomes such as learning (through, for example, tracking of staff turnover, staff absenteeism, and process quality in the provision of services, all of which have been linked to learning), the critical link that the subnational level of governance could serve between national policy planning and local implementation can be very weak or broken (Pritchett, 2015).

Subnational features that may be measurable and associated with lower-level systems functioning and ultimately local program quality include a variety of vertical links: resources and information that pass up to the national level and down to lower level systems; the roles, duties and capacities of the subnational workforce; political or

other sources of variation at the subnational level and the impact on carrying out these duties; and the capacity to process and make decisions based on data from the local or municipal level. Better understanding of this level might be particularly helpful for the "small to bigger" approach to scale. Understanding why some subnational units have strong ECD programs or policies and what it would take to expand them at a national level is a research agenda that is understudied in most country contexts. Leveraging differences in access at the subnational level can also be an analytic tool for understanding causal impacts of policies at scale (e.g., Hoynes, Schanzenbach, & Almond, 2016).

Vertical links and actions based on information from the two other levels. Links of exchange of information or resources can be understood as links up from the subnational level to the national level and links down to lower-level systems of implementation. Links up include accurate reporting of data, particularly the forms of data that may be more closely linked to children's outcomes than simple numbers of families served. National data called for in SDG Indicator 4.2.1 may spur more countries to actually gather child-level data and report it up through the subnational level to the national level. Links up may also include the passing of resources, such as revenues specific to funding ECD policies or programs that are gathered at municipal or household levels. Finally, these links may include information about variation in implementation within the subnational level (e.g. across municipalities, or across cultural, ethnic, or linguistic groups). The previously mentioned *modalidad propia* in Colombia is an example of a feedback process from an entire cultural group up to the national level to incorporate subnational definitions of quality in the national quality standards and monitoring system.

Links down include the capacity of the subnational level to monitor and improve the functions of district or municipal systems of service oversight, training, and finance. To what degree, for example, do subnational actions actually affect the functioning of district or lower-level systems? Take the example of India, a decentralized society with much responsibility at the state level for budgeting and lower-level systems functioning. The Indian administrative system at the state level sits atop a multi-layer system in the education sector of districts (ranging from several hundred to several thousand schools) to clusters to village- or neighborhood-based organizations. Links down also include the adequacy of resources passed through from the national level to the local or municipal level (if indeed there is a pass through—in some cases resources for ECD programs travel directly from the federal level to the municipal or even household level, as with many social protection policies like conditional cash transfer programs). There is a good deal of potential for "slippage" at the subnational level that can create a "leaky bucket" such that national level resources and budgeting do not reach down to the municipal and local levels or only do so inconsistently. To understand the circumstances of such slippage and examples of positive links of resources and information for implementation, detailed study of the roles and incentives affecting actors at the state, district, and cluster levels is needed. Such a study—essentially a multi-site ethnographic study of the roles, responsibilities, daily constraints, and incentives affecting the work duties of all education staff at the block and district level offices (which are respectively responsible for hundreds and thousands of government schools in India)—was recently conducted by Aiyar and

Bhattacharya (2016). Both in-depth interviews and time-use studies (conducted through a series of half-day visits to block and district staff) showed that block-level staff focused almost entirely on conveying higher-level directives to lower-level staff, and on hiring and union issues, rather than on using monitoring data from lower levels to inform higher-level decision-making. The directionality of governance was thus entirely top-down and, moreover, did not stretch down to school-management stakeholders such as parents except when engaging the local level only involved headmasters. Such a study has not been conducted in ECD in India. The at-scale Integrated Child Development Services scheme, for example, has a different administrative structure than primary government schooling.

This issue of levels and the connections or links between levels is one of the central tenets of systems analysis. Understanding the factors and agents of any system almost always begins with the organization of those elements by the level at which they are situated, but much of dynamic systems theory is concerned with understanding the relation between levels (Grobman, 2005). Systems research has explored how factors at multiple levels can function both jointly and independently, influence each other across levels, and operate in non-linear ways. The need to better understand this virtually unexamined middle level, between the more researched national and municipal levels, will be critical.

Characteristics of the subnational administrative workforce. The responsibilities of those specific ECD workers with roles that span entire subnational units (typically within subnational governance and leadership) are often the least defined, relative to either national ministry roles (e.g., early childhood leads in line ministries such as health or education) or local roles (e.g., community health worker, home visitor, early childhood teacher or caregiver). This is because national quality standards for ECD programming typically do not include specificity about what a subnational staff member responsible for an ECD program should do or what constitutes quality at that level of governance and program oversight. Without clear definitions of duties and what constitutes metrics of performance—beyond the basics of data reporting, convenings, and visits to municipal level programs—the links to local-level program quality and, ultimately, to children's outcomes may be extremely weak (Pritchett, 2015). In lower-income countries in particular, a "brain drain" may draw the higher-capacity staff regarding leadership duties up to the national level, leaving the subnational level under-resourced. And at a more basic level, the staffing at the subnational level may simply not be adequate to support the professional development and duties of lower-level staff. Consider an example where a single province-level staff member may be responsible for the monitoring of hundreds of municipal-level programs. Such a scenario could be exacerbated by the travel time it may take to visit remote or rural programs, language barriers due to language diversity across the relevant communities, or other barriers.

Measurement challenges at the subnational level include the following. First, accuracy of data passed up and down the national-subnational-local corridors should be assessed (and the overall directional pattern; cf. Aiyar and Bhattacharya's 2016 study). With inaccurate or insufficiently frequently collected data, there will be little capacity for the national system to respond to subnational data in ways that can inform policy implementation (e.g., regarding access, quality, or equity).

Second, the workforce responsibilities at the subnational level should be defined so that their assessment is possible. What constitutes the capacity to assess accurately the functioning of lower-level systems across health, education, and social protection? These include the capacity to assess the adequacy of training systems and institutions; lower-level workforce compensation, attendance, qualifications, and supports; managing the political or emergency- or conflict-related complexities that can affect local implementation; and the capacity to accurately gather and monitor finances (if this is within the purview of the subnational level) of both federal and subnational origin. Once defined, these can be integrated into national monitoring and support of subnational implementation of ECD policies and programs.

Approaches to measurement: Disaggregating subnational data and measuring subnational processes. The ability to adjust policy implementation at the local level involves subnational decision-making that could be enhanced with relevant data. For example, identifying quality gaps across municipalities, identifying coverage gaps (especially for marginalized groups) and how these differ across municipalities, and engaging political leadership at the municipality level so that local level cooperation and coordination in delivering the resources and services in multi-sectoral ECD programs is achieved all require the active gathering of data, both quantitative and qualitative (e.g., based on key informants), across municipalities within a subnational unit in a way that accurate disaggregation can occur. Even if such data are gathered, whether the subnational level of leadership can act on these data may depend on the political structure of the nation. The political party affiliation at the subnational level relative to the municipal levels, for example, may be either an enabling factor or a constraining factor in the ability to make use of data and intervene when an access, equity, or quality gap is identified at the municipal level.

Information on how data are gathered and used for policy and practice decisions at the subnational level could be gathered systematically. To our knowledge, there have been virtually no efforts to do this in the ECD sector in any country. The closest examples come from some countries where researchers have studied how data are used in organizational contexts, such as subnational or lower-level policy structures (Coburn & Turner, 2012). Such a research agenda may be useful if applied to the ECD field at the subnational level.

Municipal or Local Factors and Their Measurement

The most local level of scale that we consider in this article is that of the municipality. We defined this level as a jurisdiction or nested level that lies above the program-site level (spanning multiple sites), but below the subnational level. As with the subnational level, municipalities vary dramatically in size and population. Large cities, for example, can often operate with great independence and power, while other towns might be so small as to be virtually indistinguishable from a single program-site level. And as previously mentioned, the definition of different levels will vary significantly across different contexts.

We recognize that certain features relevant to ECD program-level implementation do not occur at a municipal or politically-defined jurisdiction level. Rather they may constitute the unit of training institution that is responsible for multiple programs but

does not coincide with the boundaries of a municipality. Similarly, district health clinics often coincide with political district definitions regarding their boundaries, but often do not. The level of governance, oversight, or training responsibility that spans multiple program sites, and the organizations and institutions at that level, are our concern here. As shorthand, we will use the word "municipal" or "local" to refer to these varied definitions below.

At the municipal level, several dimensions may constrain or enable quality implementation at the program-site level. These include features of training and monitoring systems, other workforce supports and characteristics, and municipal-level governance and budgeting.

Training and monitoring systems spanning multiple program sites. Training systems are critical for support of the workforces in ECD. They are typically of two types, pre-service and in-service. Pre-service systems are centered in training or higher-education institutions. The distribution and coverage of these institutions in a country can impact the proportion of a workforce that completes certificates, degrees, or other training programs. For example, a governance study of Cambodia's ECE programs observed that only one training institution for the state preschool program existed, and that was in the capital (Britto et al., 2014). Since that time, regional training institutions have begun to be established in certain provinces of the country.

The focus of training on the specific capacities and skills of providers that are related to child development and learning seems like a very basic principle. However, this feature of training is often not measured well in ECD systems. Consider whether a municipal or district-level trainer or inspector understands and can provide feedback regarding the quality of interactions of health workers or caregivers (however these are defined within a given cultural or national context). Can that trainer or inspector convey feedback regarding this observed skill on the part of the provider? As a field, ECD research lacks assessments of this focus (or lack thereof) in training materials and trainer skills. Without a focus on provider skills that matter for children's development and learning, again a gap occurs in the larger system between training and monitoring systems on the one hand, and child outcomes on the other (Pritchett, 2015). Some recent efforts employ online, asynchronous, video-based feedback to trainers to assess whether they are recognizing and fostering more responsive interactions between child care providers and children. Such programs may be useful in reaching areas with professional development supports in the absence of actual physical visits, assuming at least periodic (Fisher et al., 2016; Pianta et al., 2016). However, these professional development programs do not target those at the municipal level, but rather the immediate trainers, mentors, or coaches of front-line providers (who might cover up to hundreds of providers). If in most country contexts the municipal level encompasses much more than this number of providers, then we must consider the capacities of the next higher levels in the administrative hierarchy to understand how to support lower-level trainers and staff. Such intentional avenues for providing feedback to municipal-

We must consider the capacities of the next higher levels in the administrative hierarchy to understand how to support lower-level trainers and staff.

level supervisors or inspectors are virtually unstudied and unmeasured in the field of ECD. The link between monitoring and actual learning on the part of providers is a key accountability gap that must be addressed if ECD systems are to be aligned with development and learning, rather than simply access (Pritchett, 2015).

Both implementation science, with its focus on family- and provider-level dosage and experienced quality of services, and organizational or systems research, which can link such family and provider factors to monitoring and training systems, can help address local accountability and quality gaps. Given that this is the most localized level, shorter and more direct feedback loops with service providers and beneficiaries can be encouraged. This connection often underlies what makes municipalities and community-based programs sites for innovation (Hayden & Wai, 2013). Autonomy might accordingly be an important factor of this level—for example, the ability to adapt and change national policies to better suit local populations or experiment with the development of new programs and services. Autonomy at the individual level is also associated with greater workforce motivation and engagement, and thus may be important to foster (Ryan & Deci, 2010). However, balancing fidelity to the original program model with the need to allow for adaptation to local contexts continues to be a challenge.

One example of an initiative to link municipal-level administration with local program quality in ECD is a training system in India within the Integrated Child Development Services system (ICDS) that incorporates a cascade system across administrative levels between municipal/local and program site. A leading NGO in the ECD sector in India—the Centre for Learning Resources—instituted improvements in the monitoring system between the district, cluster, and local program levels such that administrators at each level taught for a time in ICDS centers, and then were trained in the monitoring of program-level quality and coaching approaches. They then brought their expertise back to their own administrative levels, implementing their new knowledge in immediate interactions with officials one level up and down in the hierarchy (Centre for Learning Resources, 2017). This approach is somewhat similar to other cascade models of training that have been evaluated in, for example, foster care systems in the United States (Chamberlain, Price, Reid, & Landsverk, 2008), but more complex in that the Indian system involves more sub-levels, even within the local/municipal level.

Other municipal-level workforce supports and characteristics. Other workforce characteristics could be placed at the national, subnational or municipal levels. Because of evidence that these characteristics vary across municipalities, we include them here.

Whether a workforce carries out its expected duties at the local level relevant to an ECD program at scale depends, beyond the skills mentioned above, on resources that include time and compensation. Time includes the ability to fit responsibilities for a specific ECD program into other duties and income generation activities of a workforce. Research on health systems found that successful coverage and impacts of HIV, AIDS, and tuberculosis treatments during the MDG era relied on task shifting, or reliance on community health workers and paraprofessionals to take on some duties that had traditionally been performed by more skilled (and scarce) nursing and medical staff. By building tasks such as medication monitoring and distribution and some diagnostic tasks into the work of community health workers, considerable improvement in disease outcomes at scale was achieved (Van Damme, Kober, & Kegels, 2008). In

contrast, recent implementation assessment of a pilot Care for Child Development parenting program in the context of rural Malawi showed that community health workers could not feasibly add this program's implementation into their many other duties (Gladstone et al., 2014). Teacher-led implementation of the "teaching to the right level" differentiation of instruction to skill level in India produced student gains during the summer, when teachers were free of school-year pressures, but not during the school year (Banerjee et al., 2016). In another area of India, once a dedicated hour was devoted to the program and local-level government officers were trained to provide mentorship and coaching, the model was more successful. In this case building in local-level systems of support and resources helped move this model further in the "small to bigger" scale process.

In ECD research on the workforce, the focus has been almost entirely on front-line providers, rather than their supervisors or those higher up at municipality-wide levels.

Beyond level of compensation, as is well documented, in low-resource contexts provider workforces that are supposed to be paid are often not paid on time or at all. These features of workforce support vary by municipality depending on aspects such as local corruption, resources for finance institutions, fluctuations in available funds at national or subnational levels, or emergencies such as disasters or armed conflict.

These workforce characteristics are typically not routinely gathered at the national level or tracked in their variation across subnational and local levels in many countries. Such data have been gathered mainly in discrete research studies. Moreover, the workforce literature has tended to focus on economic features, such as pay and incentives, accountability or psychological features, such as motivation and burnout, rather than the ECD-specific roles and their implementation within the context of other roles and responsibilities. Finally, in ECD research on the workforce, the focus has been almost entirely on front-line providers, rather than their supervisors or those higher up at municipality-wide levels.

Municipal-level governance and budgeting. Municipal-level governance and budgeting is the key to "last mile" differences in implementation that ultimately may reflect in the degree of national progress in ECD policy. In some nations, responsibility for budgeting for ECD is partly at the local level. These are due in some cases to decentralization policies (Britto et al., 2014). For example, municipal support and contributed finance from that level have been instituted to supplement federal funds in the implementation of integrated nutrition and learning interventions at the local level in the context of a decentralization process in Senegal (UNICEF Senegal Office, 2016). As at the subnational level, whether budgeting actually results in intended spending will matter for services that are actually received. In addition, the level of support and training provided for such budgeting may be important in ensuring that ECD programs are indeed supported among the various other municipal priorities. County clerks in charge of local budgeting in some countries of the six-country governance study (e.g., the commune level in Cambodia) reported little systematic training for their roles (Britto et al., 2014).

Governance factors relevant to ECD may include gender representation, given the cross-national evidence on gender differences in intra-household and other investment patterns (Behrman, 1997). In India, Pathak and Macours (2013) showed that greater representation of women on village governing councils (panchayat) was causally associated with better young child health outcomes in those villages, with suggestive evidence that increased spending priorities of the panchayat on water and sanitation were responsible. This study implies that inclusion of other under-represented or traditionally marginalized groups in local governance may improve outcomes for those particular groups.

Variation in resource levels and access at the municipal level may also affect ultimate impacts of ECD policies. One recent study examining heterogeneity in impacts of the national Head Start preschool program in the United States showed that impacts on vocabulary were larger in urban than in rural communities, while impacts on some other outcomes showed the reverse pattern (McCoy, Morris, Connors, Gomez, & Yoshikawa, 2016). Further exploration of this moderation showed that differences in access to transportation at the local level may have explained some of this variation.

Approaches to measurement and mixing methods to understand local variation. In addition to the need for disaggregation to the municipal level, which is analogous to that of disaggregation discussed previously in the subnational level section of this report, several aspects of measurement can be considered at the local or municipal level. However, municipal characteristics are typically not directly examined in multi-site evaluations, aside from their demographic differences. Quality of program site-level implementation may vary due to municipal or local workforces, as well as governance, finance, and resource differences, in addition to compositional differences.

Local governance and budgeting processes are typically not systematically tracked or measured. This level is closest to that of traditional impact evaluation and, in fact, many multi-site evaluations incorporate multiple municipalities. In the absence of more comprehensive tracking at the municipal level of ECD workforce, governance, and resource factors, multi-site evaluations could include efforts to monitor some of these factors as potential additional sources of impact variation. To the extent that these factors could then be embedded in the more general monitoring and evaluation (M&E) systems of local actors—implementing NGOs, government—there may be local improvements in data collection. Program-level monitoring checklists, typically the way in which national standards are monitored at the program level, could also include some of these factors at least as they are perceived by program directors and other stakeholders.

Mixed qualitative–quantitative methods, in addition, have been employed to explore local variation in impacts in some ECD and other child program evaluations (Huston, Duncan, & Yoshikawa, 2016; Yoshikawa, Weisner, & Lowe, 2006). For example, in social protection policies in the United States, the social-welfare offices, which represent individual jurisdictions, have been examined relative to their organizational climates, goals, and practices. Variation in these has been examined from qualitative perspectives (Meyers, Glaser, & Macdonald, 1998), and linked quantitatively to variation in child impacts in multi-site experiments (Godfrey & Yoshikawa, 2012).

Conclusion

The research agenda for "small to bigger" and "big to better" forms of scale is urgent in the context of the U.N. SDGs and simultaneous new large investments in early childhood development. With decades of strong evidence in neuroscience and program evaluation, we note the striking dearth of systems-level data that can inform the global crisis in both the provision and quality of ECD services. Systems-level supports for ECD programs range from professional development systems at municipal scale; feedback and accountability links based on accurate data regarding services, governance, and budgeting; and indicators of subnational and national planning, coordination, and budgeting across sectors. Yet virtually no research exists on any of these processes. We aimed to set a research agenda by focusing on systems-level processes that could be the focus of research as countries aim to achieve quality early childhood development, care, and education as called for in Sustainable Development Target 4.2 as well as related goals and targets relevant to ECD in the health, social protection, and other areas.

Given the magnitude of such a potential research agenda, what are feasible steps to advance research on these two types of scaling? First, we believe that building on existing data collection efforts may be a feasible initial step. At the national level, our review highlighted several constructs that are not currently collected in the major global effort on this front, the World Bank SABER ECD assessment. It may be more practical to build on this existing cross-national effort rather than suggest these elements of data collection to all countries in a diffuse way. At the local level, we suggest that multi-site impact evaluations may gather some of the local variation in governance, finance, and workforce characteristics, in addition to the typical demographics of families and children that are gathered and disaggregated at the local/municipal level. And at the subnational level, the SDGs require disaggregation by potentially marginalized groups such as by ethnicity, urban/rural variation, and disability status—these can be extended to ensure that equity at subnational levels can be not only assessed but interpreted from the perspective of ECD policy implementation.

The research capacity to conduct such systems-level measurement and analysis could also be built in the ECD field. Global networks on issues of scale can support this effort (e.g., an effort by Cooley and Linn [2014] has led to a global network of policy researchers in different areas of international development working on measuring and tracking systems-level factors related to scaling programs).

Several cross-cutting methodological issues are relevant to a systems-level research agenda. For instance, among the indicators we reviewed (cf. Table 1), some may be more tractable for quantitative measurement, and others for qualitative analysis. For example, the dynamics of the building of political will for sustainable ECD programs and policies or the motivations, constraints, and perspectives of subnational-level actors may best be analyzed qualitatively. Many other indicators may be most instructive as population and systems-level quantitative indicators. The need for both qualitative and quantitative data will require multiple and varied types of research.

Second, investments to support research on systems-level factors associated with quality implementation at national, subnational, and municipal levels are currently inadequate in most country contexts. We believe that informing ECD policy at large

scale will be improved by building systems-level factors into national data systems, such as those we highlight in this article. Therefore, many of the constructs and dimensions we discussed were phrased as elements of national data systems rather than research projects.

Third, the study of systems-level factors may clash with the culture of ECD research. Most researchers in ECD focus on individual or family-level processes. After all, the study of young children occurs in the context of the thousands of everyday interactions occurring in the first years of life between children and their parents, family, community members, and frontline service providers. Thus, the vast majority of ECD research occurs at the levels of the child, the family, and the immediate proximal settings (such as preschool, child care, household, and in some cases community or service settings). The systems-level approach that has been put forward in areas like health services (Van Damme et al., 2008), mental health services (Nadeem et al., 2014), and primary education (Pritchett, 2015) is sorely lacking in the field of ECD. Integrating this systems-level focus into training in ECD research may be needed to foster a new generation of researchers spanning the proximal contexts of human development with data on national, subnational, and municipal policy and program processes. This will require further extensions of an already multi-disciplinary research field.

Integrating this systems-level focus into training in ECD research may be needed to foster a new generation of researchers spanning the proximal contexts of human development with data on national, subnational, and municipal policy and program processes.

Finally, we note limitations of this review. We have concentrated on the national, subnational, and municipal levels of scale and implementation. This leaves the global or cross-national level of support for national ECD programs and policies missing. Relative to the national-level dimensions of scale considered here, cross-national initiatives are a yet more macro-level systems factor. Funding, information, data and measures, human capital, leadership, and other resources can be shared across nations in ways that can strengthen at-scale ECD programs within countries. The most obvious example is the role of multilateral NGOs and donor agencies in supporting cross-national coordination. Cross-national networks (of governments or civil society organizations) can also play an important role in sharing resources. Examples of current regional networks that are active in ECD include the African Regional ECD Network, the Arab Network for ECD, and the Asian Resource Network for Early Childhood (ARNEC), with global networks as well such as the new ECD Action Network (2016). These may benefit from links to an emerging global network of policy researchers addressing scale in different sectors, coordinated by Management Systems International (Cooley & Ved, 2012).

We further have not considered the methodological issues related to causally linking the systems-level factors we propose for measurement to program-level quality or child development. This requires a level of technical review that is beyond the scope of this paper. However, we can note the few examples cited here that attempted to

strengthen causal inference by, for example, leveraging geographic variation in the rollout of particular policies (e.g., of women's political reservation in India; Pathak & Macours, 2013), by leveraging variation in the rollout of services at the subnational level (Hoynes et al., 2016), or by using random assignment to examine workforce issues or program design issues (comparing impacts of the implementation of the same program by NGO versus government workforces; Bold et al., 2013; Banerjee et al., 2016). Another approach is to manipulate, within an existing service system, a key hypothesized mechanism for effectiveness. If an individual behavior or perception within a large-scale system might be key for implementation quality or impacts on children, an individual-level experiment may be relevant to policies implemented at scale (Heller, Shah, Guryan et al., 2015).

Last but not least, we have focused as much as possible on ECD-specific factors. This means we have left out more general features of quality policy implementation at the national level such as basic good governance and corruption or, at the subnational or local level, the functionality of distribution systems (roads, food, or other material resource distribution networks).

Despite these limitations, we hope that the constructs and dimensions outlined here supplement the current focus of the ECD research field on individual development and family and program-level processes with an agenda at the systems level that may inform the critical goal of quality programs and policies at national scale. Only by doing so can the promise of supporting human and societal development inherent in the field of early childhood development truly be fulfilled.

References

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research, 38,* 4-23. doi: 10.1007/s10488-010-0327-7

Aiyar, Y., & Bhattacharya, S. (2016). The post office paradox: A case study of the block level education bureaucracy. *Economic and Political Weekly, 51*(11), 61-69.

Araujo M., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *Quarterly Journal of Economics, 131,* 1415-1453. doi: 10.1093/qje/qjw016

Arbour, M., Yoshikawa, H., Atwood, S., Duran, F. R., Godoy, F., Trevino, E., & Snow, C. E. (2015). Quasi-experimental study of a learning collaborative to improve public preschool quality and children's language outcomes in Chile. *BMJ Quality and Safety, 24*(11), 727. doi: 10.1136/bmjqs-2015-lHlabstracts.11

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., Walton, M. (2016). *From proof of concept to scalable policies: Challenges and solutions, with an application (NBER working paper w22931).* Cambridge, MA: National Bureau of Economic Research. doi: 10.1257/jep.31.4.73

Behrman, J. R. (1997). Intrahousehold distribution and the family. In M.R. Rosenzweig & O. Stark (Eds.), *Handbook of Population and Family Economics: Volume 1* (pp. 125-187). Amsterdam: North-Holland. doi: 10.1016/S1574-003X(97)80021-9

Berlinski, S., & Schady, N. (2015). *The early years: Child well-being and the role of public policy.* New York: MacMillan, and Washington, DC: Inter-American Development Bank.

Bernard van Leer Foundation (2011). (Ed.). *Early learning: Lessons from scaling up.* The Hague, Netherlands: Author.

Berwick, D. M. (2003). Disseminating innovations in health care. *JAMA, 289*(15), 1969-1975. doi: 10.1001/jama.289.15.1969

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D., Fink, G., Shawar, Y., Shiffman, J., Devercelli, A., Wodon, Q., Vargas-Baron, E., Grantham-McGregor, S. (2017). Early childhood development coming of age: science through the life course. *The Lancet, 389,* 77-90. doi: 10.1016/s0140-6736(16)31389-7

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., & Sandefur, J. (2013). *Scaling up what works: Experimental evidence on external validity in Kenyan education (working paper 321).* Washington, DC: Center for Global Development. doi: 10.2139/ssrn.2241240

Britto, P.R., Yoshikawa, H., Van Ravens, J., Ponguta, L.A., Reyes, M., Oh, S.S., Dimaya, R., Nieto, A.M., & Seder, R. (2014). Strengthening systems for integrated early childhood development services. *Proceedings of the New York Academy of Sciences, 1308,* 245-255. doi: 10.1111/nyas.12365

Britto, P.R., Lye, S., Proulx, K., Yousafzai, A., Perez-Escamilla, R., Rao, N., Ip, P., Fernald, L., MacMillan, H., Hanson, M., Wachs, T., Yao, H., Yoshikawa, H., Cerezo, A., Leckman, J., Bhutta, Z. and the Early Childhood Development Interventions Review Group, for the *Lancet* Early Childhood Development Series Steering Committee (2017). Nurturing care: Science and effective interventions to promote early childhood development. *The Lancet, 389* (10064), 91-102. doi: 10.1016/S0140-6736(16)31390-3

Britto, P., Yoshikawa, H., & Boller, K. (2011). Quality of early childhood development programs and policies in global contexts. *Social Policy Reports, 25*(2), 1-31.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R.M. Lerner (Eds.), *Handbook of Child Psychology* (pp. 793-828). New York: Wiley.

Brouwers, S. A. (2017). The positive role of culture: What cross-cultural psychology has to offer to developmental aid effectiveness research. *Journal of Cross-Cultural Psychology*, [on line] doi: 0022022117723530.

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: growth curves from an early childhood educational experiment. *Developmental Psychology, 37*, 231-242. doi: 10.1037/0012-1649.37.2.231

Carr, M., & May, H. (1993). Choosing a model. *International Journal of Early Years Education, 1,* 7-22. doi: 10.1080/0966976930010302

Centre for Learning Resources (2017). *Annual report: 2016-2017.* Pune: Author. Downloaded November 3, 2017 from: http://www.clrindia.net/downloads/Annual_Report_2016_17.pdf

Chamberlain, P., Price, J., Reid, J., & Landsverk, J. (2008). Cascading implementation of a foster and kinship parent intervention. *Child Welfare, 87*(5), 27-48. doi: PMC2676450

Chaudry, A., Morrissey, T., Weiland, C., & Yoshikawa, H. (2017). *Cradle to kindergarten: A new plan to combat inequality.* New York: Russell Sage Foundation.

Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational Evaluation and Policy Analysis, 30*(3), 203-235. doi: 10.3102/0162373708321829

Coburn, C. E., & Turner, E. O. (2012). The practice of data use: An introduction. *American Journal of Education, 118*(2), 99-111. doi: 10.1086/663272

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., Ramey, S. L., Shure, M. B., & Long, B. (1993). The science of prevention: a conceptual framework and some directions for a national research program. *American Psychologist, 48,* 1013-1022.Cooley & Linn, 2014

Cooley, L., & Ved, R. (2012). *Scaling up—from vision to large-scale change: a management framework for practitioners (2nd edition).* Washington, DC: Management Systems International.

Dahlberg, G., Moss, P., & Pence, A. R. (1999). *Beyond quality in early childhood education and care: Postmodern perspectives.* New York: Psychology Press.

De Savigny, D. & T. Adam 2009. *Systems thinking for health systems strengthening.* Geneva: WHO.

Denboba, A., Hasan, A., & Wodon Q. (2015). (Eds.). *Early childhood education and development in Indonesia: An assessment of policies using SABER.* Washington, DC: World Bank.

Dodge, K. A. (2011). Context matters in child and family policy. *Child Development, 82,* 433-442. doi: 10.1111/j.1467-8624.2010.01565.x

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., Romanelli, L. H., Leaf, P. J., Greenberg, M. T. & Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools. *Advances in School Mental Health Promotion, 1,* 6-28. doi: 10.1080/1754730X.2008.9715730

Doyle, O., Fitzpatrick, N., Lovett, J., & Rawdon, C. (2015). *Early intervention and child health: Evidence from a Dublin-based randomized controlled trial.* Dublin: University College Dublin.

ECD Action Network (2016). *Technical consultation report.* New York: Author. Downloaded March 5, 2017 from: https://anecd.mawared.org/sites/default/files/ecdan_rep_6-16_f.pdf

Fisher, P. A., Frenkel, T. I., Noll, L. K., Berry, M., & Yockelson, M. (2016). Promoting healthy child development via a two-generation translational neuroscience framework. *Child Development Perspectives, 4,* 251-256. doi: 10.1111/cdep.12195

Fixsen, D., Blase, K., Naoom, S., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19,* 531-540. doi: 10.1177/1049731509335549

Ford, D. H., & Lerner, R.M. (1992). *Developmental systems theory: An integrative approach.* Thousand Oaks, CA: Sage.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M. & Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science, 344*(6187), 998-1001. doi: 10.1126/science.1251178

Gladstone, M., Maleta, K., Koenraads, M., Mirdamadi, S., Rahman, A., & Phuka, J. (2014). *Building brains in Malawi.* Paper presented at Academy Medical Science Winter Meeting, London.

Godfrey, E. B., & Yoshikawa, H. (2012). Caseworker–recipient interaction: Welfare office differences, economic trajectories, and child outcomes. *Child Development, 83,* 382-398. doi: 10.1111/j.1467-8624.2011.01697.x

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., & Himes, J. H. (1991). Nutritional supplementation, psychosocial stimulation, and mental-development of stunted children: The Jamaican study. *The Lancet, 338,* 1-5. doi: 10.1016/0140-6736(91)90001-6

Grobman, G. (2005) Complexity theory: A new way to look at organizational change. *Public Administration Quarterly, 29,* 350-382.

Guthrie, D. (2001). *Dragon in a three-piece suit: The emergence of capitalism in China.* Princeton University Press.

Hayden, J., & Wai, S. (2013). Community-based approaches to early childhood development. In P. Britto, P. Engle, & C. Super (Eds.), *Handbook of early childhood development research and its impact on global policy* (pp. 275-290). Oxford: Oxford University Press.

Heller, S., Shah, A., Guryan, J., Ludwig, J., Mullainathan, S., & Pollack, H. (2015). *Some field experiments to reduce crime and dropout in Chicago (Intitute for Policy Research Working paper).* Evanston, IL: Northwestern University. doi: 10.1093/qje/qjw033

Hoynes, H., Schanzenbach, D. W., & Almond, D. (2016). Long-run impacts of childhood access to the safety net. *The American Economic Review, 106,* 903-934. doi: 10.1257/aer.20130375

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2,* 88-110. doi: 10.1080/19345740802539325

Huston, A.C., Duncan, G.J., & Yoshikawa, H. (2016). Mixed methods contributions to understanding anti-poverty policies: Four case studies. In C. M. Hay (Ed.), *Methods that matter: Integrating mixed methods for more effective social science research.* Chicago: University of Chicago Press.

Institute for Healthcare Improvement (2007). *Whole system measures.* Cambridge, MA: Author.

Kagan, S.L., & Kauerz, K. (2012). (Eds.). *Early childhood systems: Transforming early learning.* New York: Teachers College Press.

Kroelinger, C. D., Rankin, K. M., Chambers, D. A., Roux, A. V. D., Hughes, K., & Grigorescu, V. (2014). Using the principles of complex systems thinking and implementation science to enhance maternal and child health program planning and delivery. *Maternal and Child Health Journal, 18,* 1560-1564. doi: 10.1007/s10995-014-1586-9

Kruk, M. E., Larson, E., & Twum-Danso, N. A. Y. (2016). Time for a quality revolution in global health. *The Lancet Global Health, 4,* e594-e596. doi: https://doi.org/10.1016/S2214-109X(16)30131-0

Lingard, L., Regehr, G., Orser, B., Reznick, R., Baker, G. R., Doran, D., Espin, S., Bohnen, J. & Whyte, S. (2008). Evaluation of a preoperative checklist and team briefing among surgeons, nurses, and anesthesiologists to reduce failures in communication. *Archives of Surgery, 143,* 12-17. doi:10.1001/archsurg.2007.21

Lu, C., Black, M. M., & Richter, L. M. (2016). Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level. *The Lancet Global Health, 4,* e916-e922. doi: 10.1016/S2214-109X(16)30266-2

Masten, A. (2007). Resilience in developing systems: Progress and promise as the fourth wave rises. *Development and Psychopathology, 19*(3), 921-930. doi:10.1017/S0954579407000442

McCoy, D. C., Morris, P. A., Connors, M. C., Gomez, C. J., & Yoshikawa, H. (2016). Differential effectiveness of Head Start in urban and rural communities. *Journal of Applied Developmental Psychology, 43,* 29-42. doi: 10.1016/j.appdev.2015.12.007

Metz, A., Naoom, S. F., Halle, T., & Bartley, L. (2015). An integrated stage-based framework for implementation of early childhood programs and systems. *Research Brief #2015-48*. Washington, DC: Administration for Children and Families.

Meyers, M. K., Glaser, B., & Donald, K. M. (1998). On the front lines of welfare delivery: Are workers implementing policy reforms? J*ournal of Policy Analysis and Management, 17,* 1-22.

Motta, A. C., & Yoshikawa, H. (in press). Challenges of Sustainable Development Goal 4 for a culturally diverse world: Infusing culture in ECD policy. In S. Verma & A. Petersen (Eds.), *Developmental science and the Sustainable Development Goals.* New York: Springer.

Myers, R. G. (1984). *Going to scale (report).* New York: UNICEF and Consultative Group on Early Childhood Development.

Nadeem, E., Olin, S. S., Hill,L. C., Hoagwood, K. E., & Horwitz, S. M. (2014). A literature review of learning collaboratives in mental health care: use but untested. *Psychiatric Services, 65,* 1088-1099. doi: 10.1176/appi.ps.201300229

Pathak, Y., & Macours, K. (2013). *Women's Political Reservation, Early Childhood Development and Learning in India (Young Lives working paper).* Oxford: Young Lives.

Pence, A. R., & Marfo, K. (2008). Early childhood development in Africa: Interrogating constraints of prevailing knowledge bases. *International Journal of Psychology, 43*(2), 78-87. doi: 10.1080/00207590701859143

Pianta, R., Downer, J., & Hamre, B. (2016). Quality in early education classrooms: Definitions, gaps, and systems. *The Future of Children, 26*(2), 119-137. doi: 10.1353/foc.2016.0015

Pidufala, O. (2008). *Scaling up and aid effectiveness: Annotated bibliography*. Washington, DC: Brookings Institution, Wolfensohn Center for Development.

Pritchett, L. (2015). *Creating education systems coherent for learning outcomes: Making the transition from schooling to learning.* Cambridge, MA: Center for Global Development.

Rhodes, R. A. (1997). *Understanding governance: Policy networks, governance, reflexivity and accountability.* Buckingham, UK: Open University Press.

Richter, L. M., Daelmans, B., Lombardi, J., Heymann, J., Boo, F. L., Behrman, J. R., Lu, C., Lucas, J. E., Perez-Escamilla, R., Dua, T. & Bhutta, Z. A. Stenberg, K., Gertler, P., Darmstadt, G., with the Paper 3 Working Group and the *Lancet* Early Childhood Development Series Steering Committee (2017). Investing in the foundation of sustainable development: pathways to scale up for early childhood development. *The Lancet.* doi: 10.1016/S0140-6736(16)31698-1

Robinson, J. P., Winthrop, R., & McGivney, E. (2016). Millions learning: Scaling up quality education in developing countries. Washington, DC: The Brookings Institution.

Rosero, J., and H. Oosterbeek. 2011. *Trade-offs between Different Early Childhood Interventions: Evidence from Ecuador.* Amsterdam: University of Amsterdam and Tinbergen Institute.

Ryan, R. M. & Deci, E. L. (2010). *Self-determination.* New York: Wiley.

Schindler, H. S., Fisher, P. A., & Shonkoff, J. P. (2017). From innovation to impact at scale: lessons learned from a cluster of research–community partnerships. *Child Development, 88,* 1435-1446. doi: 10.1111/cdev.12904

Serpell, R. (2011). Social responsibility as a dimension of intelligence, and as an educational goal: Insights from programmatic research in an African society. *Child Development Perspectives,* 5(2), 126-133. doi: 10.1111/j.1750-8606.2011.00167.x

Shiffman, J. & Smith, S. (2007). Generation of political priority for global health initiatives: A framework and case study of maternal mortality. *The Lancet, 370,* 1370-1379. doi: 10.1016/S0140-6736(07)61579-7

Shonkoff, J. P., & Phillips, D. A. (2000). (Eds.). *From neurons to neighborhoods: The science of early childhood development.* Washington, DC: National Academy Press.

Skocpol, T. (1995). State formation and social policy. In T. Skocpol, *Social policy in the United States: Future possibilities in historical perspective* (pp. 11-36). Princeton, NJ: Princeton University Press.

Sun, J., Rao, N. and Pearson, E. (2015). *Policies and strategies to enhance the quality of early childhood educators (Background paper for EFA Global Monitoring Report 2015).* Paris: UNESCO.

Supplee, L.H., & Metz, A. (2015). Opportunities and challenges in evidence-based policy. *Social Policy Reports of the Society for Research in Child Development, 28*(4), 1-31.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: Bradford Books / MIT Press.

Trochim, W. M., Cabrera, D. A., Milstein, B., Gallagher, R. S., & Leischow, S. J. (2006). Practical challenges of systems thinking and modeling in public health. *American Journal of Public Health, 96,* 538-546. doi: 10.2105/AJPH.2005.066001

Twum-Danso, N. A. Y., Dasoberi, I. N., Amenga-Etego, I. A., Adondiwo, A., Kanyoke, E., Boadu, R. O., Atinbire, S., Balagumyetime, P., Bagni, F., Kubio, C., Sagoe-Moses, I., & Barker, P. M. (2014). Using quality improvement methods to test and scale up a new national policy on early post-natal care in Ghana. *Health Policy and Planning, 29,* 622-632. doi: 10.1093/heapol/czt048

UNESCO Institute for Statistics (2016). *Global Alliance to Monitor Learning: Technical document.* Montreal: Author.

UNICEF, Senegal Country Office (2016). *Le développement de la petite enfance au Sénégal: Note technique.* Dakar: Author.

United Nations (2017). *Revised list of SDG indicators.* New York: Author.

Valenzuela Delpiano, P., & Cordero Vega, M. (2011). Success factors in an integrated early childhood development policy. *Early Childhood Matters, 117,* 12-16.

Van Damme, W., Kober, K., & Kegels, G. (2008). Scaling-up antiretroviral treatment in Southern African countries with human resource shortage: how will health systems adapt? *Social Science and Medicine, 66*(10), 2108-2121. doi: 10.1016/j.socscimed.2008.01.043

Van Ravens, J., & Aggio, C. (2008). *Expanding early childhood care and education: How much does it cost?* The Hague: Bernard van Leer Foundation.

Vargas-Barón. (2015). *Policies on Early Childhood Care and Education: Their evolution and some impacts.* Paris: UNESCO.

Vegas, E., & Ganimian, A. (2011). *Education diagnostics: A tool to identify binding constraints of education systems.* Washington, DC: The World Bank.

Von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications.* New York: George Braziller.

Waddington, C. H. (1957). *The strategy of the genes: A discussion of some aspects of theoretical biology.* New York: Macmillan.

Wiek, A., Ness, B., Schweizer-Ries, P., Brand, F. S., & Farioli, F. (2012). From complex systems analysis to transformational change: a comparative appraisal of sustainability science projects. *Sustainability Science, 7,* 5-24. doi: 10.1007/s11625-011-0148-y

Woodhead, M., & Streuli, N. (2013). Early education for all: Is there a role for the private sector? In P. Britto, P. Engle, & C. Super (Eds.), *Handbook of early childhood development research and its impact on global policy* (pp. 308-328). Oxford: Oxford University Press.

World Bank (2016). *SABER early childhood development (SABER brief).* Washington, DC: Author.

Wuermli, A. J., Hempel, K., Aber, J. L., & Lundberg, M. (2012). Policies to protect and promote young people's development during crisis. In M. Lundberg & A. J. Wuermli (Eds.), *Children and youth in crisis: Protecting and promoting human development in times of economic shocks.* Washington, DC: World Bank.

Yoshikawa, H., & Hsueh, J. (2001). Child development and public policy: Toward a dynamic systems perspective. *Child Development, 72,* 1887-1903. doi: 10.1111/1467-8624.00384

Yoshikawa, H., Rosman, E. A., & Hsueh, J. (2002). Resolving paradoxical criteria for the expansion and replication of early childhood care and education programs. *Early Childhood Research Quarterly, 17,* 3-27. doi: 10.1016/S0885-2006(02)00129-1

Yoshikawa, H., Weisner, T. S., & Lowe, E. (2006). (Eds.). *Making it work: Low-wage employment, family life and child development.* New York: Russell Sage.

Young, M. E. (2003). *From early childhood development to human development: Investing in our children's future.* Washington, DC: The World Bank.

Yousafzai, A. K., & Aboud, F. (2014). Review of implementation processes for integrated nutrition and psychosocial stimulation interventions. *Annals of the New York Academy of Sciences, 1308*(1), 33-45. doi: 10.1111/nyas.12313

Author Bios

Hirokazu Yoshikawa is the Courtney Sale Ross Professor of Globalization and Education at NYU Steinhardt and a University Professor at NYU. He co-directs the NYU Global TIES for Children Center. He is a community and developmental psychologist who studies the effects of public policies and programs related to immigration, early childhood, and poverty reduction on children's development. He conducts research in the United States and in low- and middle-income countries. His recent books include *Cradle to Kindergarten: A New Plan to Combat Inequality* (with Ajay Chaudry, Taryn Morrissey, and Christina Weiland, 2017) and *Immigrants Raising Citizens: Undocumented Parents and Their Young Children* (2011). He serves on the Board of Trustees of the Russell Sage Foundation, and on the Advisory Boards for the Open Society Foundations Early Childhood Program and the UNESCO Global Education Monitoring Report. In 2014, he was elected to the National Academy of Education, and in 2018 he was elected to the American Academy of Political and Social Science as the James S. Coleman Fellow. He directs the evaluation research for the MacArthur Foundation 100&Change funded partnership between Sesame Workshop and the International Rescue Committee. He obtained his PhD in clinical psychology from NYU.

Alice Wuermli is a post-doctoral Research Scientist and Associate Director for Program at Global TIES for Children, New York University. Alice is interested in how human ecosystems can be changed to improve development, health, and well-being from conception to young adulthood in low- and middle-income countries. Particular interests pertain to how an individual's neurobiology interacts with contexts; the mechanisms through which interventions effect change; how stress physiology mediates/moderates intervention effectiveness; how changes in neurobiological markers inform questions of dose-response; and advancing tools to meaningfully and validly measure contexts, processes, and outcomes. In her capacity at Global TIES, Alice is working on multiple research projects in several low-, middle-, and conflict-affected countries, and across a variety of topics including early language development in pre-school children; home visiting, parenting, and early development; and implementation research to support scale and quality in early childhood development services. Alice holds a PhD in Human Development from the University of California, Davis; a MA in International Development/Development Economics from American University in Washington, DC; and a BA in International Affairs and Governance from the University of St. Gallen, in Switzerland.

Abbie Raikes, PhD., MPH, is an associate professor at the College of Public Health, University of Nebraska Medical Center, and a Fellow at the Buffett Early Childhood Institute. Dr. Raikes' recent work has focused on improving early childhood programs and policies in low- and middle-income countries. Her research background also includes a strong focus on young children's social/emotional development. Through her leadership of the Measuring Early Learning and Quality Outcomes project, Abbie's work has supported low- and middle-income countries in adopting effective measurement systems for early childhood development. Previously, Abbie contributed to early childhood policy development in several countries as a program specialist for the United Nations Education, Science and Culture Organization (UNESCO) in Paris. Abbie was a senior program officer at the Bill & Melinda Gates Foundation, and has advised several organizations on early childhood development and education. Abbie has served on several boards, including the Nebraska Early Childhood Collaborative, the Consultative Group on Early Childhood Development, and the Gretchen Swanson Center for Nutrition.

Sharon Kim is a doctoral student in Psychology and Social Intervention at NYU. She holds a B.A. in English Literature and Music from the University of Michigan as well as an M.A. in International Educational Development from Columbia University. Prior to beginning her doctoral studies, she spent three years working with Dr. Edward Seidman on the development of the Teacher Instructional Practices and Processes System (TIPPS), a systematic behavioral observation tool that aims to understand the quality of the classroom environment in low-, middle-income, and conflict-affected countries. Her research interests since beginning her doctoral career have included 1) understanding the instructional and socio-emotional factors that influence a quality learning environment and how those factors manifest in various cultural contexts; 2) the development and refinement of instruments to measure such factors; and 3) the effective integration of relevant findings within the respective cultural contexts through more rigorous implementation research. Sharon has served as a research consultant for various INGOs and research organizations including the World Bank, FHI 360, and the International Rescue Committee. Sharon currently serves as project Director for TIPPS.

Sarah Kabay is a PhD candidate in International Development Education at NYU Steinhardt's School of Education, Culture, and Human Development. Her research focuses on primary school, early childhood education, and how education fits into greater development and social change frameworks. Current research projects include the cost analysis of Boston Public Schools' public pre-K program, the impact assessment of a school management intervention in Ugandan primary schools, and a qualitative investigation into the systems-level factors that influence educational interventions in Niger. Before beginning her doctoral program, Sarah lived in Uganda for five years, working to conduct randomized controlled trials of development interventions with the organization Innovations for Poverty Action. Much of her work continues to focus on Uganda, where she collaborates with schools, NGOs, local government officials, and other education stakeholders to conduct research and disseminate findings. She is a 2017 Spencer Dissertation Fellow, a Research Affiliate of the Global TIES for Children Research Center, and she serves on the board of the organization Yspaniola and as technical adviser for Elevate: Partners in Education.

Social Policy Report is a quarterly publication of the Society for Research in Child Development. The *Report* provides a forum for scholarly reviews and discussions of developmental research and its implications for the policies affecting children. Copyright of the articles published in the SPR is maintained by SRCD. Statements appearing in the SPR are the views of the author(s) and do not imply endorsement by the editors or by SRCD.

Purpose

The Social Policy Report (SPR) is a quarterly publication of the Society for Research in Child Development (SRCD). Its purpose is twofold: (1) to provide policymakers with comprehensive, non-partisan reviews of research findings on topics of current national interest, and (2) to inform the SRCD membership about current policy issues relating to children and about the state of relevant research.

Content

The SPR provides a forum for scholarly reviews and discussions of developmental research and its implications for policies affecting children. Topics are drawn from a variety of disciplines and cover a wide range of issues that affect child and family development through the lens of social policy, such as health care, parenting practices, and education policies. SRCD recognizes that few policy issues are noncontroversial and that authors may well have a "point of view," but the SPR is not intended to be a vehicle for authors to advocate particular positions on policies. Presentations should be balanced, accurate, and inclusive. The publication nonetheless includes the disclaimer that the views expressed do not necessarily reflect those of the SRCD or the editors.

Procedures for Submission and Manuscript Preparation

Articles originate from a variety of sources. Some are solicited by the editorial board, while others are proposed by the authors. Authors interested in submitting a manuscript are urged to propose timely topics to the lead editor via email. Topic proposals should take the form of an extended abstract (approximately 2 pages) that outlines the topic and scope of the proposed report. Manuscripts vary in length ranging from 30 to 45 pages of text (approximately 8,000 to 12,000 words), not including references and figures. The manuscript should be double-spaced throughout with 12-point font and should adhere to APA guidelines. Manuscript submission should include text, abstract, references, and a brief biographical statement for each of the authors and should be sent as a .doc, .docx, or .rtf file.