Published in final edited form as:

Prev Sci. 2016 October ; 17(7): 794–805. doi:10.1007/s11121-015-0611-6.

Generative Mechanisms in Early Childhood Interventions: A Confirmatory Research Framework for Prevention

Arthur J. Reynolds and Suh-Ruu Ou

University of Minnesota

Arthur J. Reynolds: ajr@umn.edu; Suh-Ruu Ou: sou@umn.edu

Abstract

This article reviews methodological and analytic approaches, and impact evidence for understanding the mechanisms of effects of early childhood interventions, including delinquency and violence prevention. Illustrations from longitudinal studies of preschool preventive interventions are provided. We restrict our attention to preventive interventions for children from birth to age 5, including evidence from the Chicago Longitudinal Study (CLS), which investigates the impact of an established school-based early childhood intervention. Frameworks and evidence will be organized according to the 5-Hypothesis Model (5HM), which postulates that a variety of early childhood interventions impact later well-being through the promotion of cognitive and scholastic advantages, motivational advantages, social adjustment, family support behaviors, and school supports. Recommendations are made for advancing confirmatory approaches for identifying the most effective prevention programs using identification of generative mechanisms as a major methodological criterion.

Keywords

Generative mechanisms; longitudinal studies; early childhood; preventive interventions; health and well-being

As knowledge has grown that social and educational interventions for young people can prevent the development of problematic behaviors and promote well-being, attention has shifted to the generative mechanisms or paths through which interventions impact well-being over the life course (MacKinnon, 2008; Reynolds & Ou, 2011). The identification, measurement, and assessment of generative mechanisms broaden the scope and relevance of evaluation research for violence prevention and health promotion. Knowledge of key processes that initiate and sustain effects strengthens causal explanation in effectiveness studies and in dissemination and expansion efforts to ensure reproducibility.

This paper reviews conceptual frameworks, methodological approaches, and impact evidence for identifying and understanding the mechanisms of effects of early childhood interventions on well-being, including delinquency and violence prevention. Illustrations from longitudinal studies of preschool preventive interventions are provided. We restrict our

attention to preventive interventions for children from birth to age 5. Evidence from the Chicago Longitudinal Study (CLS), which traces the effects of the Child-Parent Center (CPC) preschool program, is highlighted along with impact studies of the Perry and Abecedarian early interventions. Frameworks and evidence will be organized according to the 5-Hypothesis Model (5HM), which postulates that a variety of early childhood interventions impact later well-being through the promotion of cognitive and scholastic advantages, motivational advantages, social adjustment, family support behaviors, and school supports. Recommendations are made for advancing confirmatory approaches for identifying the most effective prevention programs using identification of generative mechanisms as a major criterion.

We cover longitudinal evidence for the impact of early childhood intervention programs in the first five years of life. Although the outcomes of violence and delinquency prevention are included, we adopt the broader term of well-being to recognize the co-occurrence of health-compromising and competence behaviors and that preventive interventions for young children are designed to impact many social and education outcomes. Interventions that promote school competence and performance may impact delinquency and crime and visa versa. These must be documented. In this report, we include delinquency and crime, school achievement and attainment, and child maltreatment as primary outcomes impacted by leading longitudinal studies early intervention. For others, see Farrington and Welsh (2007), Camilli, Vargas, Ryan, and Barnett (2010), and Manning, Homel, and Smith (2010).

Mechanisms in Prevention and Behavioral Science

The identification and understanding of mechanisms are central to scientific explanation. As reflected in philosophy of science (Popper, 1956) and theories of human development (Farrington & Welsh, 2007; O'Connell, Boat, & Warner, 2009), mechanisms enhance causal understanding and provide a framework for describing key elements and processes of behavioral change. A mechanism can generally be defined as a testable process of change involving one or more influences at any level of behavior. Mechanisms may vary from a single mediator that accounts for an observed effect to complex systems involving a chain of effects through many mediators. Within the epistemological tradition, Bunge (1997) defined a mechanism as "a process in a concrete system, such that it is capable of bringing about or preventing some change in the system as a whole or in one of its subsystems" (p. 414). We use this broad systems perspective in discussing the relevance of mechanisms in early childhood prevention programs.

Although mechanisms that account for intervention effects must be hypothesized, a major goal of prevention science is to empirically validate and corroborate hypothesized mechanisms. Only empirically supported mechanisms can strengthen causal inferences and understanding necessary for valid and generalizable knowledge.

Although many terms have been used to define the concept (e.g., mediator, causal mechanism, pathway), we use generative mechanism as the operational term for describing how participation in intervention affects later well-being. A major advantage of this term is that it directly focuses on the nature of the observed differences in outcomes between groups

(i.e., generative) without the clear presumption of causality implied by the term causal mechanism. It also is suggestive of a process of influences, possibly multiple processes, that the terms mediator or mediating mechanism do not convey.

Terms such as snowballing and cascade effects, while denoting a complex process, are terms that have broad meaning. They also suggest a singular process that may not reflect how change from interventions affects outcomes. Below we describe the relevance and use of generative mechanisms for documenting the effects of prevention programs.

Contributions of Identifying Generative Mechanisms

The methodological importance of mechanisms is acknowledged in the Society of Prevention Science's evidence standards (Flay et al., 2005; Gottfredson et al., 2015). A key requirement for effectiveness studies is that "a key theory of causal mechanisms be stated" and appropriately assessed in impact. This importance is reinforced in the newly released standards (Gottfredson et al.). Knowledge about generative mechanisms enhances methodology and design of preventive interventions in three major ways. First, its documentation enhances the capacity for causal inferences. Since the length of time between the end of the program and outcome assessment can be a decade or more, documenting the process of change is central for understanding the impacts of preventive interventions (Susser, 1973; O'Connell, Boat, & Warner, 2009). Even in well-controlled studies, there are many intervening events and experiences that challenge the capacity to infer causality (Chen, 1990; MacKinnon, 2008). If the identified mechanisms lead to a coherent explanation of the main-effect relation, causal inference is enhanced.

Confidence that an intervention enhances well-being or reduces the prevalence of economically-consequential problem behaviors such as violence and crime is critical in program dissemination and policy decisions, especially when alternative uses of funds are possible. Increased levels of confidence about intervention effects and their sources can also help broaden support for intervention strategies that may be necessary before funding or program expansion occurs. The valid measurement of core and alternative mechanisms requires comprehensive planning and resource allocation in study design.

Second, documentation of mechanisms of effects is important for program design and modification. Empirically supported paths that are identified can contribute to a variety of improvement efforts. For example, if family support behavior is identified as an important mechanism of effects in violence prevention, greater attention to parental involvement and the provision of family services would be one improvement strategy, whether or not it was connected to child maltreatment or other complex linkages. The lack of presumed linkage to parental involvement would suggest problems in implementation or with the program theory. This knowledge can also be used to design new programs or strengthen existing elements. In addition to family services, if socio-emotional factors or peer relations are sources of long-term effects, instructional activities could be better aligned to increase the likelihood of affecting these processes. A finding that developed abilities is a key source of effects would suggest a different alignment of activities such as language and literacy instruction. More

generally, the identification of mechanisms can help establish key principles of effectiveness in the design, modification, and validation of programs to promote health and well-being.

Finally, identifying and understanding generative mechanisms can increase or help clarify generalizability and external validity of findings. To the extent that findings across studies share a common mechanism (or within subgroups of a particular study), program replication and expansion in different contexts would be more likely to be successful. Knowledge of key processes that initiate and sustain effects may also be used to help ensure reproducibility. For example, the post-program school and family contexts may directly influence the strength of longer-term effects on delinquency or violence prevention. This information can also explain why interventions did not yield consistent or enduring effects. Empirical support for an identified theory of change may also inform the development of other interventions or strategies. For example, an intervention that is found to reduce school dropout by enhancing school climate could also be broadened to prevent delinquency using the same mediator or “active ingredient” of school climate. Alternatively, the consistency of evidence comparing two interventions for preventing delinquency but through different mechanisms could inform program priority for expansion and scale up. Home visitation to improve parenting skills and early childhood socio-emotional skill-building are two such alternatives. The improving the quality of evidence on mechanisms, however, would directly inform decision making.

Although increased attention to generative mechanisms has advanced understanding about preventive interventions, few program registries such as the National Registry of Effective Programs and Practices and Model Program Guide of the Office of Juvenile Justice and Delinquency Prevention) assess or account for the identification of mechanisms from the perspective of internal or external validity. Only Blueprints for Violence Prevention states that it addresses mediators of effects (i.e., mechanisms) but it is only a supplemental criterion and no information is provided as to how much weight, if any, is given to identification of mediators in the selection of model and promising programs. We attribute this gap to the higher priority given to research design and main-effect estimation than to causal explanation and generalizability. The relative absence of procedures and frameworks for assessing mechanisms in program impact research also hinders their use as core validity criteria.

A Confirmatory Program Evaluation Approach

To facilitate the analysis of generative mechanisms, confirmatory program evaluation (CPE) was developed (Reynolds, 1998, 2005). CPE is a theory-driven methodology for investigating the effects of social and educational programs. In a theory-driven impact evaluation, the explicit theory of the program is highlighted to establish an a priori model of how the program is expected to exert its influence. Causal uncertainty is reduced through an examination of the empirical pattern of findings against the expectations inherent in the program. Contrast this perspective with a purely method-driven approach, in which causal uncertainty is reduced through control exercised during the research design phase, or the statistical modeling approach, whereby control is exercised during the data analysis phase via statistical adjustment. These approaches, of course, can be complementary for the strengthening inferences. Confirmatory evaluation attempts to strengthen causal inference

through systematic investigation of the nature of the relationship between program participation and outcome. Special emphasis is given to testing generative mechanisms.

Confirmatory evaluation is an impact evaluation in which hypotheses about the sources of effectiveness are tested based on the program theory. The primary objective is to facilitate causal inference about the link between program participation and measured outcomes. Given its focus on impact, confirmatory evaluation is distinct from evaluability assessment (i.e., logic model assessment) and implementation evaluation. Although it can be applied to experimental, quasi-experimental, and observations studies, it strengthens inferences for the latter two designs most. This is because validity threats are typically better controlled in experiments than in the other designs. Confirmatory evaluation also benefits most from longitudinal investigations with multiple post-program assessments. Consequently, the pattern of findings across domains can be fully documented and linked to the theoretical predictions.

Three key questions are addressed: (1) Is program participation independently and consistently associated with key outcomes? (2) Do the estimated effects vary by background characteristics, such as child and family attributes, or by program components? (3) What are the processes or pathways through which participation leads to effectiveness in the short term and over time? Given their emphasis on multivariate prediction, theory-driven evaluations are more confirmatory than other evaluation approaches. Implementation assessment would inform all three questions, but would primarily be addressed in the first question. The power or capacity to influence outcomes is dependent on good-quality implementation.

Based on the program theory, for example, an evaluator can explicate or test the following: (a) the size of the program effect, (b) the program outcomes that yield the largest and the smallest effects, (c) the consistency of effects across subgroups, models, and analyses, (d) the generative mechanisms through which the estimated effects are manifested, and (e) the factors that may influence selection into the program and implementation quality. Analyses of theory-driven evaluations are often conducted through traditional regression procedures but often include structural modeling and path analysis. Because of the primary focus on understanding causal mechanisms of change, findings from confirmatory evaluations also can enhance generalizability of findings in that identified mechanisms can be applied to broader classes of intervention and different social contexts.

Table 1 shows the major steps involved with conducting a confirmatory evaluation. The key is to describe as completely as possible the program theory in terms of (a) the background factors and experiences that may influence participation, implementation level, and outcome measures; (b) measures of participation, such as timing, duration, intensity, and quality; (c) mechanisms through which program participation leads to short-, intermediate-, and long-term outcomes; and (d) primary and secondary outcomes, including their timing and measurement. Note that tests of alternative program theories are desirable, especially if the program theory is complex or not well established.

Confirmatory evaluation uses six empirically verifiable criteria that strengthen the validity of findings (Anderson et al., 1980; Chen, 1990; Reynolds, 2005; Susser, 1973). Meeting these criteria increases the likelihood that the link between program participation and outcomes is causal, and they help probe analyses of theory-relevant questions. These are, in order of importance from lowest to highest, temporality of program exposure, size or magnitude of association, gradient or dose-response effects, specificity, consistency, and coherence. These criteria were used in the Surgeon General's Report establishing the causal link between smoking and lung cancer (U. S. Department of Health, Education, & Welfare, 1964). Because of the key importance of coherence, which is primary established through the identification of generative mechanisms, we further describe its relevance.

At the highest level of causal interpretation, the coherence criterion assesses whether the pattern of findings provide a convincing story about the effects of the program. Coherence is defined as the extent to which the theoretical predictions of the program are consistent with the observed findings in terms of size, strength, and consistency. The main investigation concerns empirical support for the hypothesized generative mechanisms. If support is found, both coherence and causal inference are strengthened. The extent to which the preventive effects of an early intervention on later outcomes are explained by the presumed mechanism (e.g., early academic competence) would provide corroborating evidence.

One landmark study (Consortium for Longitudinal Studies, 1983) demonstrated this path of early cognitive-scholastic advantage to adolescent social competence through a process of cumulative advantages in school experiences. Other studies have implemented a mechanisms approach to strengthen validity. In the Parents as Teachers home visiting program from birth to age 3, for example, Zigler, Pfannenstiel, and Setiz (2008) found that impacts on third-grade achievement for a large state sample were explained by improvements in school readiness and theory-related changes in home literacy and post-program preschool enrollment.

The availability of a wide range of data is beneficial for establishing coherence. Ethnographic and qualitative data on program implementation are complementary to quantitative approaches. The coherence criterion is especially relevant to prevention programs with extensive longitudinal assessments. Relatively little attention in the evaluation field has been devoted to generative mechanisms, or more broadly, the individual, program, and contextual conditions that help maintain or enhance effectiveness.

Phases of Research on Generative Mechanisms in Early Childhood Intervention

Research over five decades shows that educational enrichment in the early years of life promotes health and well-being in many domains from school entry to adulthood (Camilli et al., 2010; Karoly, Kilburn, & Cannon, 2005). Long-term effects of early intervention on delinquency and criminal behavior, educational attainment, mental health, and socioeconomic status are especially significant (Campbell, Ramey, Pungello, & Ramey, 2002; Consortium for Longitudinal Studies, 1983; Schweinhart, Barnes, & Weikart, 1993,

Schweinhart et al., 2005; Reynolds, Temple, Robertson, & Mann, 2001; Reynolds et al., 2007).

While there is now a critical mass of evidence in support of long-term effects, the generative mechanisms that account for these findings are not well understood. Comprehensive models have rarely been tested and differential effects across a range of outcomes and programs are unknown (Cuhna & Heckman, 2008; Reynolds & Temple, 2008).

Three phases of research on generative mechanisms have occurred. In the earliest studies representing the first phase, the cognitive advantage hypothesis was found to be the initiator and primary mediator of long-term effects on school achievement, educational attainment, and economic well-being (Berrueta-Clement, Schweinhart, Barnett, Epstein, & Weikart, 1984; Consortium for Longitudinal Studies, 1983). Cognitive advantage, as assessed by IQ test scores, was the sole mediator of long-term effects in the Carolina Abecedarian Project (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001). Other mediators were not directly assessed.

The second phase of research added family and motivational factors to better document the complexity of influences (Barnett, Young, & Schweinhart, 1998; Reynolds, Ou, & Topitzes, 2004; Schweinhart et al., 1993). For example, Seitz et al. (1983) found that cognitive advantage was partly a function of increased motivation in accounting for the impact of a Head Start program on school achievement. Barnett et al. (1998) found that the initial cognitive advantage in the Perry preschool led to later advantages through motivation and social adjustment ultimately accounting for effects on educational attainment. Generally, these studies demonstrated support for the cognitive advantage hypothesis but less support for other mechanisms.

In recognition of the complex array of factors both during and after program participation that account for long-term effects, the third phase of research has emphasized investigating a comprehensive set of child, family, and school-related mechanisms within a framework of alternative models. This led to the development of the 5-hypothesis model of intervention effects (5HM; Reynolds, 2000).

Five-Hypothesis Model of Effects for Early Interventions

Derived from ecological (Bronfenbrenner, 1989) and risk/protection (Rutter & Rutter, 1993) theories of human development as well as the accumulated research on early intervention over four decades, the 5HM posits that the effects of intervention are explained by indicators of five general paths of influence: cognitive advantage, motivational advantage, social adjustment, family support behavior, and school support behavior. A major tenet of this comprehensive approach is that contribution of any one hypothesis can be accurately assessed only in conjunction with the others. School quality, for example, has been found to account for links between preschool participation and later school performance (Currie & Thomas, 2000) but has not been included in studies that assess other hypotheses. Within the framework of 5HM, investigation of alternative models also is essential to determine the combination of hypothesized paths that fit the data best. The model has been found to have

adequate fit with observed data and explain substantial percentages of direct effects of intervention (Reynolds & Ou, 2011; Reynolds et al., 2004).

Figure 1 shows the five hypothesized paths through which early intervention affects well-being into adulthood. Because the major purpose of early intervention is to promote school success and social adjustment, enduring effects into adulthood are frequently investigated.

Derived from the literature on the effects of early intervention (Consortium for Longitudinal Studies, 1983; Ou & Reynolds, 2006; Reynolds & Temple, 2008; Schweinhart et al., 1993), the theory-based five hypotheses provide a foundation for developing the knowledge base on how early childhood programs lead to longer term effects, and the personal and environmental conditions that may limit success. As shown, the effects of early intervention may be transmitted through (a) developed cognitive and scholastic abilities (cognitive advantage hypothesis), (b) social development and adjustment (social adjustment hypothesis), (c) parents' behavior with or on behalf of children (family support hypothesis), (d) children's motivation or self-efficacy (motivational advantage hypothesis), and (e) the quality of the school environments that children experience post-program (school support hypotheses). Consistent with ecological and risk/protection theories, these mechanisms are mutually reinforcing.

They account for observed effects through a process of cumulative advantage. Early cognitive and scholastic impacts enhance the transition to school and promote social adjustment, which create further advantages that "snowball" to greater adult well-being. For example, early skill advantages can strengthen the quality of family and school environments and reduce the risks of frequent mobility and need for remedial education (e.g., special education) and social services (e.g., child welfare; Reynolds & Ou, 2011).

Note that the hypotheses must be independently associated with both program and outcome measures to be valid mechanisms of long-term effects. That is, the arrows on each side of the intervening constructs in Figure 1 must be significant. Of course, the hypotheses could work in combination. For example, program participation may affect social competence through early developed abilities and family support experiences or through a combination of early developed abilities, family support, and school support factors. Empirically supported hypotheses may vary by program and sample characteristics.

Analyses of Life-Course Well-Being of Early Intervention

In this section, we summarize evidence from studies of early childhood interventions that have tracked participants through adolescence and into adulthood. First, findings are summarized from the Chicago Longitudinal Study of the Child-Parent Center (CPC) program. This is followed by a synthesis of evidence from other interventions on delinquency and crime, education, and child maltreatment. Each of the programs reviewed are identified in the six registries of effective programs described above. Next, generative mechanisms of the effects of the CPC program are addressed using 5HM. Finally, the generalization of the model is tested across three studies (CPC, Perry, and Abecedarian) of effects on educational attainment.

Effects of CPC Intervention

The CPC program was established in 1967 through funding from Title I of the Elementary and Secondary Education Act of 1965 to counteract through educational enrichment the negative effects of poverty to promote school success (Sullivan, 1971). Throughout its history, the 25 CPCs were located in the 20 poorest city neighborhoods. CPC children attended schools, for example, in which 67% of children resided in low-income families compared to 42% for all children (Reynolds, 2000; Temple, Reynolds, & Miedel, 2000).

The CPC program provides comprehensive educational and family support to economically disadvantaged families. The core principles are its school-based structure, emphasis on literacy, and strengthening the family-school relationship. The program has a strong literacy curriculum, in conjunction with approaches to develop other social and emotional development. Similar to grade school, the CPC program focuses on an array of activities, including individualized learning, small-group activities, and frequent teacher feedback. To maximize individual learning opportunities, preschool class sizes are small, and each classroom has a teacher's aide in addition to a regular classroom teacher. The average teacher to child ratio is 1 to 8. This small class size enables an individualized approach to cognitive and social development. The combined literacy and child-focused program within a developmentally appropriate ecological framework is central for promoting child well-being. The program is also unique in that preschool to third grade services are included (Reynolds & Temple, 2008).

Research on the effectiveness of the CPC program is based on many cohorts of graduates and a diverse set of studies (Reynolds, 2000). The primary one is the Chicago Longitudinal Study (CLS, 2005), an on-going prospective study of a complete cohort of 989 children born in 1979–1980 who graduated from kindergarten in 1986 in 20 CPCs and a matched comparison group of 550 children of the same age who attended publicly funded full-day kindergarten in 5 randomly selected schools and in schools serving the CPC program. The comparison group enrolled in the usual early childhood programs for low-income children including Head Start. The groups were equivalent on child and family characteristics at the beginning of the study.

The performance of CPC preschool participants consistently exceeded that of the comparison group at the beginning of kindergarten and into adulthood for many indicators of well-being, including cognitive skills at school entry, school achievement, delinquency and crime, and substance misuse (see Reynolds et al., 2001, 2007, 2011a). Although effect sizes varied by outcome, most were above 0.20 standard deviations, which translate to substantial social benefits.

The pattern of findings suggests that the initial effects of the program on cognitive skills ($ES = .59$ SD) at age 5 as measured by the composite scale of the Iowa Tests of Basic Skills, contributed to a cumulative advantage on later well-being (Reynolds et al., 2004, 2009, 2011b). Program-related reductions in need for special education placement ($ES = -.45$ SD) and grade retention ($ES = -.37$ SD) as well as lower rates of later delinquency and crime are indicative of significant economic benefits. For example, by age 24 the preschool group had

a 22% lower rate of felony arrest (court records) than the comparison group (16.5% vs 21.1%, respectively; Reynolds & Ou, 2011; Reynolds et al., 2011b).

Generative Mechanisms to Felony Arrest

We tested the 5HM to account for the observed main effects of CPC preschool on felony arrests by age 24. The evaluation compared a complete cohort of over 900 children participating in the program for 1 or 2 years beginning at age 3 to a matched comparison group participating in the usual early education interventions. Measured by official court records from age 18–24, felony arrests are for more severe criminal behavior in which the identification of generative mechanisms are particularly important. Based on a theoretically-grounded a priori sequence of measures of the five hypotheses (Reynolds & Ou, 2011), the full model substantially accounted for the impact of CPC preschool to felony arrest. Not only did the five hypotheses explain 79% of the main effect of preschool, but the model fit the observed data by multiple criteria (Adjusted Goodness of Fit Index = .917; RMSEA = .068; RMR = .033) and better than other nested models (Reynolds & Ou, 2011).

As shown in Figure 2, the mechanisms leading from preschool to felony arrest were diverse. Two intervening mechanisms accounted for the main effect of preschool and showed the largest indirect effects. One was a family support indicator–abuse/neglect path, and the other was school support path–number of school moves. Preschool participation was associated with lower incidence of child abuse and neglect ($b = -.13$), which in turn linked to felony arrest ($b = .12$). Preschool participation was associated with fewer school moves ($b = -.13$), which in turn linked to felony arrest in the expected direction ($b = .09$). The thicker paths in Figure 2 show the direct mediation of abuse/neglect and mobility without contributions from other mechanisms.

The cognitive advantage hypothesis contributed to the effects circuitously. Preschool participation was associated with an immediate cognitive advantage ($b = .36$) that led to a higher rate of attendance in high-quality schools ($b = .24$), higher ITBS reading scores ($b = .07$), higher classroom adjustment ($b = .38$), a lower rate of grade retention or special education placement ($b = -.15$), a lower incidence of child abuse and neglect ($b = -.16$), and to a higher level of parent involvement in school ($b = .19$). Cognitive advantage also contributed to arrest through other mediators except for school commitment. The major predictors of felony arrest were ITBS reading ($b = -.13$), high school completion ($b = -.37$), juvenile arrest ($b = .22$), and abuse and neglect ($b = .12$). The 5HM also accounted for the main effects of intervention on depressive symptoms (89% of main effect; Adjusted Goodness of Fit Index = .914) and occupational skills (58% of main effect; Adjusted Goodness of Fit Index = .914; Reynolds & Ou, 2011).

Summary of Main Effects for Other Early Interventions

Although there is a substantial literature on the effects of early childhood intervention, few studies have tracked participants through adolescence and into adulthood on delinquency, crime, and educational attainment. Studies of five programs are highlighted below: Perry Preschool (Schweinhart et al., 2005), Child-Parent Centers (CPC; Reynolds et al., 2007), Abecedarian Project (Campbell et al., 2002), Nurse-Family Partnership (NFP; Olds et al.,

1997), and Head Start of the Panel Study of Income Dynamics (PSID; Garces et al., 2002). These programs are included in the registries of effectiveness described earlier. We summarize findings for four life-course outcomes.

Crime and educational attainment—As shown in Table 2, both the Perry and CPC programs demonstrated significant program effects on crime. These effects were large. Participation in the Perry program was associated with a 40% decrease in arrests by age 19, whereas participation in the Chicago program was associated with at 33% reduction in juvenile petitions by age 18. These effects endured into adulthood for both programs. Black PSID Head Start participants also demonstrated lower rates of adult arrest but not whites. Note that Head Start participation was retrospectively reported in this study although a fixed-effects approach strengthened inferences. The lack of crime prevention benefits in Abecedarian may be due to the low base rates of crime in North Carolina, or to the relative absence of family program services.

Participation in CPC, Perry, and Abecedarian programs was linked to significantly higher levels of education well into adulthood. This was not found for the Nurse-Family Partnership (Eckenrode et al., 2010) or for PSID Head Start black participants. For the latter, White participants did show higher rates of high school completion (Currie & Thomas, 2000; Garces, Currie, & Thomas, 2002). Overall, these findings show that the programs enhanced participants' long-run well-being despite their major differences in social context and educational approaches. These impacts led to high economic returns (Reynolds & Temple, 2008).

Child maltreatment—For the outcome of child maltreatment, a strong predictor of delinquency and crime, relatively few early interventions have assessed impacts on official child abuse and neglect. The two most investigated are NFP and CPC. Olds et al. (1997) found that rates of official child maltreatment by age 15 in the Elmira trial were 24% for the full intervention group and 32% for the control group. This is a 25% reduction over the controls. Among the highest risk sample (single mothers in poverty), the preventive effect on maltreatment substantiation was greater (19% vs 42%, respectively). In the CPC program, the rate of official child maltreatment by age 18 was 43% lower in the preschool group than in the comparison group (9.9% vs 17.4%, respectively). Findings were similar for receipt of child welfare services (6.1% vs 11.3%). Effects were larger for child neglect than for abuse. See Reynolds, Mathieson, and Topitzes (2009) for additional findings on this and other studies.

Cognitive and social development—Finally, Camilli et al. (2010) conducted a meta-analysis of 123 studies of early childhood interventions reported from 1960 to 2000. Notably, the studies represent a wide range of quality and generally do not have the service intensity associated with the most effective model programs reported in the registries. Average effect sizes in standard deviations across all reported outcomes were .231 for cognitive outcomes, .156 for social development, and .137 for school progress indicators such as need for special education and grade retention. These significant direct effects are predictive of longer-term outcomes, such as antisocial behavior, delinquency, and school dropout. Crime prevention is a major source of economic benefits (Cohen, 2005).

Generative Mechanisms across Three Early Interventions

For a final set of results, the generalizability of the 5HM was assessed for CPC, Perry, and Abecedarian interventions, which demonstrated positive effects on educational attainment at age 21. Using a matched set of indicators for each of the five hypotheses, a structural model was estimated following a sequence similar to that shown in Figure 2 (Reynolds et al., 2010). We highlight two sets of findings below.

Each of the three programs demonstrated significant effects on one or more proximal outcomes measured from ages 5 to 8. Coefficients are standardized and adjust for measurement error, gender, and mother's education. The impact on cognitive skills at age 5 was strongest and most consistent, followed by teacher ratings of motivation. Only Perry and CPC programs directly affected classroom adjustment. Impacts on parent involvement were specific to the CPC program. Such findings are a necessary condition for establishing generative mechanisms to the later distal outcome of years of education.

Results of the full 5HM indicated that that indicators of the hypotheses accounted for 58% of the main effect on educational attainment for Perry, 100% for Abecedarian, and 60% for CPC. Common paths of influence across the studies were as follows: program participation to cognitive skills, cognitive skills to motivation, motivation to need for remedial education, and remedial education to adolescent achievement. Not surprisingly, the juvenile arrest was the strongest and most predictor of educational attainment (standardized $b = -.47$ for Abecedarian, $-.42$ for CPC, and $-.43$ for Perry). Model fit, based on multiple criteria, was highest for CPC followed by Abecedarian, and then Perry (Reynolds et al., 2010).

Generally, cognitive skills and achievement were more influential generative mechanisms for Perry and Abecedarian than CPC. School quality was more influential for CPC. Given limited measurement of school quality and mobility, the school support hypotheses could not be fully estimated for Perry and Abecedarian. Findings support the value of 5HM in identifying generative mechanisms to well-being. Examination of 5HM for other outcomes and programs is warranted.

Discussion

The implications of this analysis for the practice of research can be organized according to the unique contributions of measuring and identifying generative mechanisms discussed at the beginning of the article: (a) strengthening causal inferences for policy decisions, (b) inform program design, and (c) increasing generalizability and external validity. First, given the support for the identified generative mechanisms, confidence that preschool programs like those reviewed above to promote youth and adult well-being is substantially increased. Impacts on delinquency, criminal behavior, and child maltreatment are suggestive of significant cost savings to society (Levin & Belfield, 2007; Lynch, 2007). The theory of early childhood interventions such as the CPC program is that early enrichment of language and literacy through center-based education within a family supportive environment strengthens school success leading to greater well-being. One implication is that increasing access to preschool and other early interventions deserve high priority. The amount of evidence on the positive and enduring effects of high-quality preschool is large. There is not

only a critical mass of evidence from long-term cost-benefit analyses (Károly et al., 2005; Schweinhart et al., 2005), but increasing evidence that large-scale programs can improve school readiness and achievement, which predict delinquency and crime prevention (Farrington & Welsh, 2007).

Moreover, compared to the evidence on interventions for older children, evidence for preschool intervention is strong. Extensive longitudinal studies of other social programs are rare and those that have been conducted have not investigated processes of impact as thoroughly as early intervention (O'Connell et al., 2009). Findings show the benefits of high-quality programs and provide a framework for documenting and understanding long-term effects and cost-effectiveness for other programs. They strengthen confidence in the economic benefits of CPC and other similar interventions (Levin & Belfield, 2007; Reynolds et al., 2011a; Reynolds & Temple, 2008). A major share of the benefits, ranging from \$7–12 per dollar invested, are from school success, educational attainment, and crime prevention, of which generative mechanisms have been corroborated (Figure 2). Crime savings are nearly half of these benefits.

Second, generative mechanisms are a key source of knowledge for strengthening programs and sustaining their effects. Given the importance of entering kindergarten and the early grades with good literacy and social skills, improving the quality of preschool programs and increasing their length and intensity would be expected to strengthen the paths leading to well-being. As a public-school program, for example, teachers in the CPCs have bachelor's degrees with certificates in early childhood. Participation begins at age 3, and kindergarten is on site. The philosophy of the education component is to learn basic skills. These features are key to demonstrated benefits. Likewise, greater opportunities for parent involvement in early interventions also would be expected to strengthen the contribution of the family support hypothesis. Unlike most other programs, each CPC has a staffed resource room for coordinating activities in and outside the centers. The provision of comprehensive services is a hallmark of the evidence-based early interventions we reviewed. Such services are likely to broaden paths of influence necessary for sustained effects.

Strengthening the quality of elementary schools is likely to promote the transmission of long-term effects on violence prevention, especially for children at risk. Structural changes may include reduced class size (Finn, Gerber, Achilles, & Boyd-Zaharias, 2001), preschool to third grade programs that provide comprehensive instructional and family support services (Takanishi & Kauerz, 2008; Reynolds & Temple, 2008), and whole-school reforms such as Schools of the 21st Century (Zigler & Styfco, 1993) and the School Development Program (Comer, Ben-Avie, Haynes, & Joyner, 1999). Interventions to promote social emotional learning and school commitment (O'Connell et al., 2009) may also help strengthen learning gains as would services to reduce the detrimental effects of mobility (Titus, 2007), child maltreatment (Reynolds et al., 2009), and grade retention (Temple et al., 2000), which are key mechanisms.

Finally, study findings provide further support for the generalizability of the identified processes from preschool to adult well-being. In the CLS, key elements of the model have been found to be consistent between adolescence and adulthood and for different outcomes.

For example, the cognitive advantage, family support, and school support hypotheses account for the main effects of intervention for achievement, delinquency, remedial education, consumer skills (Reynolds, 2000), educational attainment (Ou, 2005; Reynolds et al. 2004), and adult well-being. Further support for key processes of the model has been demonstrated in the Abecedarian Project, Perry Preschool, and the Consortium for Longitudinal Studies (1983). However, generative mechanisms have not been investigated for the Nurse-Family Partnership.

Moreover, other interventions that impact these processes would be expected to contribute to enhanced well-being. These could be independent or complementary of preschool. For example, interventions that prevent child maltreatment may have longer-term effects on health and well-being through impacting juvenile delinquency, school achievement, and need for remedial education. Similarly, programs and strategies to increase early classroom adjustment, parental involvement, and school achievement as well as reduce involvement in the juvenile justice system would be likely to be effective through elements of the five hypothesis model. Combined with other studies of paths of influence, the cognitive advantage hypothesis is a consistent explanation of long-term effects. More generally, early cognitive and scholastic advantages lead to social and motivational gains that culminate in enhanced health and well-being. The generalizability of this and related processes will be critical in establishing the reproducibility of effects in other programs and contexts.

Finally, as a further recommendation, our analysis indicates that generative mechanisms should be a major criterion in assessing the validity of evidence in prevention programs and in selecting those that are exemplary or deserving of broader dissemination. Although the perceived importance of mechanisms is well-documented (Flay et al., 2005), their empirical assessment remains low. Our paper provides a comprehensive framework for increasing and improving their use. This heightened priority should lead to increased resources to their measurement and assessment in designing prevention studies. Traditionally, the field of evaluation research has emphasized internal validity over external validity in assessing programs. Most evaluation resources and funding are allocated to this focus. A unique strength of generative mechanisms is that their identification strengthens both internal and external validity. One limitation is that the data requirements for assessing mechanisms are extensive, including the collection and analysis of longitudinal data over many years and for many proximal and distal outcomes. However, the absence of empirical support for a mechanism does not invalidate findings on main effects. Its presence can strengthen empirical support and provide a level of confidence that would be difficult to achieve in other ways.

Acknowledgements

The authors are at the Human Capital Research Collaborative, University of Minnesota, 301 19th Avenue South, Minneapolis, MN 55455. Preparation of this article was supported by Grant R01 HD034294 from the National Institute of Child Health and Human Development.

References

Anderson, S.; Auquier, A., et al. Statistical methods for comparative studies: Techniques for bias reduction. New York: Wiley; 1980.

- Barnett, WS.; Young, JW.; Schweinhart, LJ. How preschool education influences long-term cognitive development and school success: A causal model. In: Barnett, W.; Boocock, S., editors. *Early care and education for children in poverty: promises, programs, and long-term results*. Albany: SUNY; 1998. p. 167-184.
- Berrueta-Clement, JR.; Schweinhart, LJ.; Barnett, WS.; Epstein, AS.; Weikart, DP. *Changed lives: The effects of the Perry Preschool Program on youths through age 19*. Ypsilanti, MI: High/Scope; 1984.
- Bronfenbrenner U. Ecological systems theory. *Annals of Child Development*. 1989; 6:187–249.
- Bunge M. Mechanism and explanation. *Philosophy of the Social Sciences*. 1997; 27:410–455.
- Camilli G, Vargas S, Ryan S, Barnett WS. Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*. 2010; 112(3):579–620.
- Campbell FA, Pungello EP, Miller-Johnson S, Burchinal M, Ramey CT. The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology*. 2001; 37(2):231–242. [PubMed: 11269391]
- Campbell FA, Ramey CT, Pungello EP, Sparling J, Miller-Johnson S. Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*. 2002; 6(1): 42–57.
- Chen, HT. *Theory-driven evaluations*. Newbury Park, CA: Sage; 1990.
- Chicago Longitudinal Study. *User's guide*. Minneapolis: University of Minnesota; 2005.
- Cohen, MA. *Costs of crime and justice*. New York: Routledge; 2005.
- Comer, JP.; Ben-Avie, M.; Haynes, NM.; Joyner, E., editors. *Child by child: The Comer process for change in education*. New York: Teachers College Press; 1999.
- Consortium for Longitudinal Studies. *As the twig is bent...lasting effects of preschool programs*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1983.
- Cunha F, Heckman JJ. Formulating, identifying, and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*. 2008; 43(4):738–782.
- Currie J, Thomas D. School quality and the longer-term effects of Head Start. *The Journal of Human Resources*. 2000; 35(4):755–774.
- Eckenrode J, Campa M, et al. Long-term effects of prenatal and infancy nurse home visitation on the life course of youths: 19-year follow-up of a randomized trial. *Archives of Pediatrics and Adolescent Medicine*. 2010; 164:9–16. [PubMed: 20048236]
- Farrington, DP.; Welsh, BC. *Saving children from a life of crime: Early risk factors and effective interventions*. New York: Oxford University Press; 2007.
- Flay BR, Biglan A, Boruch RF, et al. Standards of evidence: Criteria for efficacy, effectiveness, and dissemination. *Prevention Science*. 2005; 6(3):151–175. [PubMed: 16365954]
- Finn JD, Gerber SB, Achilles CM, Boyd-Zaharias J. The enduring effects of small classes. *Teachers College Record*. 2001; 103(2):145–183.
- Garces E, Currie J, Thomas D. Longer-term effects of Head Start. *American Economic Review*. 2002; 92:999–1013.
- Gottfredson DC, Cook TD, Gardner FEM, Gorman-Smith D, Howe GW, et al. Standards of evidence for efficacy, effectiveness, and scale-up research in *Prevention Science: Next generation*. *Prevention Science*. 2015 Forthcoming.
- Karoly, LA.; Kilburn, MR.; Cannon, JS. *Early childhood interventions: Proven results, future promise*. Santa Monica, CA: Rand Corporation; 2005.
- Levin, HM.; Belfield, C., editors. *The price we pay: Economic and social consequences of inadequate education*. Washington, DC: Brookings Institution Press; 2007.
- Lynch, RG. *Enriching children, enriching the nation: Public investment in high-quality prekindergarten*. Washington, DC: Economic Policy Institute; 2007.
- MacKinnon, DP. *Introduction to statistical mediation analysis*. New York: LEA; 2008.
- Manning M, Homel R, Smith C. A meta-analysis of the effects of early developmental prevention programs in at-risk populations on non-health outcomes in adolescence. *Children and Youth Services Review*. 2010; 32:506–519.

- O'Connell, ME.; Boat, T.; Warner, KE., editors. Preventing mental, emotional and behavioral disorders among young people: Progress and possibilities. Washington: National Academy Press; 2009. National Research Council.
- Olds D, Eckenrode J, Henderson C, Kitzman H, Powers J, et al. Long-term effects of home visitation on maternal life course and child abuse and neglect: Fifteen year follow-up of a randomized trial. *JAMA*. 1997; 278:637–643. [PubMed: 9272895]
- Ou S. Pathways of Long-term Effects of an Early Intervention Program on Educational Attainment: Findings from the Chicago Longitudinal Study. *Journal of Applied Developmental Psychology*. 2005; 26(5):578–611.
- Popper, K. The nature of scientific discovery. New York: Basic Books; 1956.
- Reynolds AJ. Confirmatory program evaluation: A method for strengthening causal inference. *American Journal of Evaluation*. 1998; 17:21–35.
- Reynolds, AJ. Success in early intervention: The Chicago Child-Parent Centers. Lincoln: University of Nebraska Press; 2000.
- Reynolds AJ. Confirmatory program evaluation: Applications to early childhood interventions. *Teachers College Record*. 2005; 107(10):2401–2425.
- Reynolds, AJ.; Englund; Ou, S.; Schweinhart, LJ.; Campbell, FA. Paths of effects of preschool participation to educational attainment: A 3-study analysis. In: Reynolds, AJ., et al., editors. Cost-effective early childhood programs in the first decade: A human capital integration. New York: Cambridge University Press; 2010. p. 415-452.
- Reynolds AJ, Mathieson LC, Topitzes JW. Do early childhood interventions prevent child maltreatment? A review of research. *Child Maltreatment*. 2009; 14:182–206. [PubMed: 19240245]
- Reynolds AJ, Ou S, Topitzes JD. Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago Child-Parent Centers. *Child Development*. 2004; 75(5):1299–1328. [PubMed: 15369516]
- Reynolds AJ, Ou S. Paths of effects from preschool to adult well-being: A confirmatory analysis of the Child-Parent Center Program. *Child Development*. 2011; 82(2):555–582. [PubMed: 21410923]
- Reynolds AJ, Temple JA. Cost-effective early childhood development programs from preschool to third grade. *Annual Review of Clinical Psychology*. 2008; 4:109–139.
- Reynolds, AJ.; Temple, JA. Economic returns of investments in preschool education. In: Zigler, E.; Gilliam, W.; Jones, S., editors. A vision for universal preschool education. New York: Cambridge University Press; 2006. p. 37-68.
- Reynolds AJ, Temple JA, Ou S, Robertson DL, Mersky JP, et al. Effects of a school-based, early childhood intervention on adult health and well-being: A 19-year follow up of low-income families. *Archives of Pediatrics & Adolescent Medicine*. 2007; 161(8):730–739. [PubMed: 17679653]
- Reynolds AJ, Temple JA, Roberson DL, Mann EA. Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-Year follow-up of low-income children in public schools. *JAMA*. 2001; 285:2339–2346. [PubMed: 11343481]
- Reynolds AJ, Temple JA, White B, Ou S, Robertson DL. Age-26 cost-benefit analysis of the Child-Parent Center early education program. *Child Development*. 2011a; 82:782–804.
- Reynolds AJ, Temple JA, Ou S, et al. School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science*. 2011b; 333:360–364. [PubMed: 21659565]
- Rutter, M.; Rutter, M. Developing minds: Challenge and continuity across the life span. New York: Basic Books; 1993.
- Schweinhart, LJ.; Barnes, HV.; Weikart, DP. Significant benefits: The High/Scope Perry Preschool study through age 27. Ypsilanti, MI: High/Scope; 1993.
- Schweinhart, LJ.; Montie, J.; Xiang, Z.; Barnett, WS., et al. Lifetime effects: The High/Scope Perry Preschool study through age 40. Ypsilanti, MI: High/Scope; 2005.
- Seitz, V.; Apfel, NH.; Rosenbaum, LK.; Zigler, E. Long-term effects of projects Head Start and Follow Through: The New Haven Project. In: Consortium for Longitudinal Studies., editor. As the twig is bent: Lasting effects of preschool programs. Hillsdale, NJ: Erlbaum; 1983. p. 299-332.
- Sullivan, LM. Let us not underestimate the children. Glenview, IL: Scott, Foreman; 1971.

- Susser, M. Causal thinking in the health sciences: Concepts and strategies of epidemiology. New York: Oxford University Press; 1973.
- Takanishi R, Kauerz K. PK inclusion: Getting serious about a P-16 education system. *Phi Delta Kappan*. 2008 Mar; 89(7):480–487.
- Temple JA, Reynolds AJ, Miedel WT. Can early intervention prevent high school dropout? Evidence from the Chicago Child-Parent Centers. *Urban Education*. 2000; 35:31–56.
- Titus D. Strategies and resources for enhancing the achievement of mobile students. *NASSP Bulletin*. 2007; 91(1):81–97.
- U.S. Department of Health, Education, and Welfare. Smoking and health: Report of the advisory committee to the Surgeon General of the Public Health Service. Washington, DC: Author; 1964. (Public Health Service Publication No. 1103).
- Zigler E, Pfannenstiel JC, Seitz V. The Parents as Teachers Program and school success: A replication and extension. *Journal of Primary Prevention*. 2008; 29:103–120. [PubMed: 18404381]
- Zigler, E.; Styfco, SJ. Head Start and beyond: A national plan for extended childhood intervention. New Haven, CT: Yale University Press; 1993.

Figure 1.
Five-Hypothesis Model of Paths from Early Intervention to Adult Well-Being

Figure 2.

LISREL model for any felony arrest by age 24, coefficients are standardized and adjusted for measurement errors.

Note. Covariates included gender and mother's education. Bolded paths represent strongest indirect effects.

Table 1**Key Steps in Confirmatory Program Evaluation**

Step	Evaluation Activity
1.	Specify program theory and processes that are expected to affect outcomes.
2.	Identify and measure outcomes and their timing over the short, intermediate, and long term. Predict outcomes that will index the largest and smallest effects of participation.
3.	Collect or obtain data on hypothesized mediators and mechanisms of the program theory and key background variables necessary to identify program impact.
4.	Estimate main effects of program participation for the total group and relevant subgroups. Investigate the temporality, size, gradient, specificity, consistency, and coherence of program effects.
5.	If main effects are detected, test hypothesized causal mechanisms of the program theory to account for estimated effects. Test alternative mediators and hypotheses. If main effects are not detected, analyze mediators to determine the possible explanations. Also assess the validity of program theory and the quality of program implementation.
6.	Interpret the pattern of findings to facilitate generalization and knowledge transfer. Use the six criteria for strengthening causal inference to guide interpretations.
7.	Identify formative uses of findings for program improvement, including modifications to program theory, program implementation, and analysis of program effects.

Note. Adapted from Reynolds (1998, 2005).

Table 2**Summary of Effects of Select Early Childhood Intervention on Well-Being**

Program	Sample Size	Prevention type	Ages of service	Rate of arrests		Years of education/HS	
				Prog	Comp	Prog	Comp
Perry Preschool	123	Selective	3–5				
Age 19				15.0%	25.0%	–10.0%	11.7
Age 27				28.0%	35.0%	–7.0%	11.9
Age 40				31.0%	48.0%	–17.0%	
Child-Parent Centers	1539	Selective	3–5				
Age 18	1404			16.9%	25.1%	–8.2% *	11.3
Age 24	1418			16.5%	21.1%	–4.6% *	12.0
Age 28	1418			19.3%	24.6%	–4.6%	12.2
Abecedarian	111	Selective	0–5	44.9%	41.2%	+3.7%	12.2
Nurse-Family Part.	400	Selective	0–2	21.3%	37.4%	–16.1% *	70.6%
Head Start (PSID)	3255	Selective	3–5				
Total	1742			4.7%	10.0%	–5.3% *	80.9%
Blacks	706			N/A	N/A	–11.6% *	N/A
White	1036			N/A	N/A	+12.2%	N/A

Note. Rates of arrest are cumulative. PSID = Panel Study of Income Dynamics. Except for Head Start, data were Prospective and arrest data were measured by administrative records. Educational attainment was assessed by self reports except for the Child-Parent Centers, which included school records and self-reports. Data for Abecedarian (Campbell et al., 2001, 2002), Nurse-Family Partnership (NFP; Eckenrode et al., 2010; Olds et al., 1997), and Head Start (Garces et al., 2002) were assessed at age 21, 19, and 24, respectively. NFP data are from Elmira, New York trial. Education attainment is high school graduation by age 19. Child-Parent Center years of education was measured by age 21 (first entry). CPC rates of arrest at age 24 and 28 are felony arrests (see Reynolds et al., 2007, 2011a, 2011b). Years of education/HS = Highest grade in school completed/High school graduation or completion.

*
p < .05