



Balancing Rigor with Complexity in Understanding the Impacts of Child Maltreatment Prevention Programs

Jon Korfmacher¹

Published online: 17 December 2019
© Society for Prevention Research 2019

Approximately 25 years ago, research teams from over 15 institutions came together to form a consortium to study the first wave of Early Head Start (EHS) programs in a randomized trial that came to be known rather prosaically as the Early Head Start Research and Evaluation Project (EHSREP). Although Mathematica Policy Research (MPR) had the primary contract for overseeing the randomization and national data collection, each of the research teams had formed a local research partnership with an EHS program in their area, and these programs from across the country made up most of the sites of the randomized trial. Each research team worked with MPR to handle data collection at their site and, in addition, proposed and conducted their own local study.

Over the years, the local research teams met with project officers and with MPR staff to plan and problem-solve (and argue about) the project at regularly scheduled meetings in Washington, DC. In addition, working groups of researchers with similar evaluation interests were formed to plan studies using common measures or constructs. Studies on fathers (Boller et al. 2006), children with disabilities (Peterson et al. 2004), parent engagement (Korfmacher et al. 2008), and child maltreatment (Green et al. 2014) followed. Many of the analyses conducted within workgroups focused on aspects of program implementation. A common refrain in discussions was a recognition of the double-edged sword of the randomized trial.

On the one hand, a large-scale evaluation with an adequate comparison group and random assignment was seen as essential for assessing the impact of a program that was receiving a significant financial investment by the federal government. This should be of no surprise to the readers of this journal. On the other hand, a randomized trial assumed a uniform treatment variable (i.e., Early Head Start services) that the

researchers all knew was not so uniform, given the extreme diversity across sites in service delivery, in program culture, and in the characteristics of the families served. There was also the not-so-small issue that Early Head Start was brand new, and the program sites were still in a formative phase, feeling their way forward and not quite sure of what they would become. Thus, the challenge was, as it always is with these sorts of trials, whether the analytic design could take into account the diversity of experiences engaged by the different kinds of families who receive different services at different program sites in determining the value of the program. In other words, the real question of interest was not “Does Early Head Start work?” but “For whom does Early Head Start best work and under what contexts?”

As Eddy and colleagues (2019) remind us in one of the four papers in this special section, this is not a new question, having been articulated (albeit in a somewhat different form) by Paul over 50 years ago (1967). Psychotherapy research, as another example, has examined “aptitude x treatment” interactions to try to uncover what treatment conditions work best for what kinds of clients (Snow 1991). This desire to better understand differential response is in many ways our proverbial “Golden Ticket.” And yet, today, we are not much closer to answering this fundamental question—what works best for whom and when—than we were in the early days of Early Head Start.

The good news is that a focus on implementation has become much more accepted in early childhood prevention research (see Schindler et al. 2019, for a related review of early childhood education research). It is now commonly acknowledged that one cannot focus merely on outcomes for children and families without understanding how to get to these outcomes. The so-called black box is increasingly opaque. This is seen not just in research in the USA but in research internationally as well (e.g., Yousafzai et al. 2018).

The not-as-good news is that much of the research on implementation is still largely descriptive, documenting the variation that is often seen among and within families and using correlational designs to try to examine meaningful patterns in

✉ Jon Korfmacher
jkorfmacher@erikson.edu

¹ Erikson Institute, 451 N. LaSalle St., Chicago, IL 60654, USA

that variation, often in post hoc analyses once the participant sample has been finalized. Prevention researchers, however, are beginning to turn their attention to precision-based approaches (e.g., August and Gewitz 2019; Supplee et al. 2018), which provide a pathway to a more systematic and a priori examination of targeted intervention approaches. An example of this is the Home Visiting Applied Research Collaborative (HARC),¹ a collaborative agreement between HRSA and the Johns Hopkins Bloomberg School of Public Health, and its mission of introducing precision-based approaches into early childhood home visiting (see Supplee and Duggan 2019, for more information).

My goal in this commentary is to use the four papers in this special section to highlight both the challenges and the potential for research to help us understand what really works. My discussion will cross three broad themes, all with implications for a precision-based approach: (1) conducting research in the “real world,” (2) individualizing and targeting approaches to different families, and (3) establishing and maintaining partnerships with programs and other key stakeholders.

Real-World Research

All intervention research is essentially applied, but efficacy trials try to control more of the wild variability that can occur when an intervention is roaming free in the community, fettering services within the confines of such things as recruitment protocols and prescriptive randomization. Each of the four papers in this section reports findings from a randomized trial with a comparison group, but in all cases, the investigators had to cede some level of control to the programs that were attempting to balance servicing the needs of their community while adhering to a randomization scheme that could appear to be denying services to some of the recruited families who needed it most.

Green and colleagues (2018), for example, in their study of the statewide Healthy Families Oregon home visiting program note their strong suspicion that the research design lowered program staff recruitment efforts, as evidenced by the high number of eligible participants who never received a single home visit. Home visitors were given the task of enrolling families into the program after randomization, with 40% unable to be found or refusing services after initially expressing some interest pre-randomization. Home visitors likely felt little incentive to go after hard to engage families, especially since there were so many other families who needed services. This can wreak havoc on an intent-to-treat analysis; the authors acknowledged this and included propensity analyses to address this challenge. The difficulty of enrolling families in services also may change the composition of the treatment

group. If harder to engage families simply stay out of the service pool altogether, how representative are the remaining families to those that the statewide home visiting program typically serves?

LeCroy and Lopez (2018) examined a Healthy Families program in a different state. As the authors note, when working with programs already established and operating within a community, families (even ones designated to the program group through a random selection lottery) can only enter a program when an opening is available in a caseload. Any staff turnover can further slow both recruitment and the assignment of families to services. Although this might be business-as-usual for programs in high need areas, long delays in recruitment are a considerable challenge to time-pressured randomized trials that have specified funding end dates. The elongated window of service delivery also increases the chances that the nature of the intervention itself may shift between earlier and later entering families.

The other two reports in this special section—an examination of a multicomponent Relief Nursery program (Eddy et al. 2019), and a study of First Steps (O’Neill et al. 2018), a much more circumscribed post-partum service—also had to accommodate to the nature of the programs operating within their community contexts. The Relief Nursery provided so many possible program components to participants that it raises the legitimate question of how well the “average” response truly represents the experiences of families and whether the program can be adequately compared to another treatment condition. Significantly, the program was also allowed some “safe” enrollment slots where families could receive intervention services if the perceived need was intense enough, eliminating the possibility that they might be assigned to the comparison group. This concession by the evaluation team was necessary for the program to accept the overall randomization plan, although in reality only two families were affected (the Oregon Healthy Families study also allowed this option for a small percentage of families). In comparison to the Relief Nursery, First Steps had a much more limited service menu, with only brief contacts with a hospital social worker. First Steps, however, also showed considerable variability across sites in the extent and content of services provided to families. For example, the two First Steps sites differed dramatically from each other in how much the specialist attempted to reach families for optional follow-up contact.

All four evaluations showed some level of positive impact of the program on child and/or family outcomes, but it seems safe to say that these impacts are relatively modest, as is true of most early childhood prevention services (e.g., Sam-Miller et al. 2018). All four of these services may have demonstrated stronger impacts if the researchers had more say in establishing the parameters of the intervention delivery that they studied. In other words, if they could more closely control the process by which families were recruited and assigned

¹ See www.hvresearch.org.

to services, as well as limit the specific service menu offered to families, it is *possible* that the treatment effects would be stronger. But what is the trade-off? How well would these trials then replicate the issues and circumstances of programs operating within their community?

In medicine, there has been a movement promoting *pragmatic* trials—a term used for designs that examine intervention effectiveness within broad routine clinical practice settings (Patsopoulos 2011). Although this term has been rarely invoked for child maltreatment prevention services research (see Ondersma et al. 2017, as an exception), much of what exists in the literature falls into the parameters of pragmatic trials. It is not revolutionary, for example, to engage in a randomized home visiting trial where services are headquartered in a community service agency, delivered in neighborhoods, and focused on real world applicability. This is what happens in much current home visiting research.

The vital lessons from pragmatic medicine trials, however, lie in the specific articulation of how these evaluations differ from the prescriptive designs of traditional explanatory studies that are well-controlled and conducted in carefully designed settings. For example, the Pragmascope (Tosh et al. 2011) presents a visual snapshot of how off-center a particular trial might be on ten different explanatory-pragmatic dimensions. Such dimensions include recruitment, setting, flexibility in service delivery, and data analysis, all of which can be seen in play in the studies in this special section. Patsopoulos (2011) notes a particular challenge of pragmatic designs is that a large sample size is needed to address the heterogeneity inherent in the study, which makes these trials more expensive to conduct. For evaluations existing in the context of reduced service funding and staffing problems, as was the case with the Relief Nursery, pragmatic designs may ironically be less practical to conduct.

Individualization of Services

All four of the programs evaluated in these studies require some level of tailoring services to the specific needs of families. This is most obvious in the Relief Nursery, where a menu of services is offered to families who can choose what they feel would be most helpful to them. But this individualization was seen in the other programs as well. The Healthy Families home visiting model similarly expects programs to develop specific family goals and can place families on different visitation schedules based on their progress and engagement in services. Although First Steps is the most prescriptive model, it also showed variation in content covered across six different areas and in whether or not parents received any follow-up content at all.

As noted earlier, most examinations of differential response have relied on post hoc examinations of subgroups to

understand moderators of effects. This is also a key feature of pragmatic trials. As Patsopoulos (2011) notes, however, the expectation is that these moderators should then be incorporated into a future efficacy trial, thereby transforming a post hoc analysis into an a priori approach. There is much less evidence of this design pathway being rigorously utilized in prevention research. Those studies that do attempt to shape interventions by targeting based on moderator variables often show mixed results (see August and Gewitz 2019).

An alternative, more dynamic approach to understanding the impact of different services is an adaptive-sequential design. This allows alteration of services to families based on their reaction to the intervention during the course of an intervention trial. In theory, this maximizes the value of the intervention for different types of families based on their real-time response to what is being offered. Guides are now available for how techniques such as sequential multiple assignment randomized trials (SMARTs; Lei et al. 2012), the multiphase optimization strategy (MOST; Guastaferrero and Collins 2019), or Bayesian approaches (Kaplan 2019) can be employed to systematically measure dynamic approaches with appropriate comparison and statistical rigor.

One can argue that the prevention programs discussed in these articles are already engaged in adaptive designs of their own making. But where does the decision-making for this tailoring of services reside? Typically, it rests with the provider—modifying dosage (e.g., when to visit), content (e.g., what is covered in a given session), or process (e.g., checking in by phone or text, changing the level of social engagement to match the needs of the family). These modifications are done based on practice judgment.

On the one hand, we value this practice judgment. Much associative research has suggested that the quality of the provider's alliance with the family is a central component of a family's willingness to stay and how much they get out of a program (Korfmacher et al. 2008; Marsh et al. 2012; Munns et al. 2016). On the other hand, there is a much we do not know about how effective this tailoring actually is and what help providers need to most effectively tailor their approach. How do providers, for example, parse the distinction between what a client *needs* and what a client *wants* when individualizing services? An example of this was provided by the study of the First Steps program. As the authors noted, their measurement of actual time spent covering different topics allowed comparing the self-reported topics in which mothers were interested with the time actually spent on those topics. Mothers did not report much interest in the topic of infant crying, even though this was the topic that the First Steps providers spent the greatest amount of time talking about with families. In turn, infant crying is where mothers showed the greatest knowledge gains, despite their lack of professed interest.

How can a systematic approach to measuring and comparing dynamic alteration of services, based by necessity on a limited set of data, adequately parallel a complex and possibly changing set of practice judgements by a home visitor working with families? Another way of framing the question is how can we support home visitors in these kinds of decisions, so that the responsibility and pressure of how to approach the individual needs of family do not rest solely on them and their own judgment? To do this, while still respecting the need for providers of services to have a say in these kinds of decisions, will require closer coordination and collaboration between programs and researchers.

Partnerships

One of the key features of a precision-based approach to prevention research is a focus on partnerships. Researchers may be careful observers of prevention services, but there is only so much an outside evaluator can truly understand about a program service. Engaging key stakeholders, especially program staff members, in the design and implementation of research is critical for prevention research to maintain relevance in practice and policy. It is necessary to help in identifying active ingredients, in identifying contexts or conditions in which interventions will or will not work, and in ensuring that the methods used are relevant and logistically practical (HARC 2019). As noted earlier, some of the trade-offs that the investigators of the evaluations in this special section had to negotiate with program staff created potential challenges in analyses and interpretation, but they were necessary for the program services to be delivered as realistically as possible in their community settings.

Different groups have established continuums of participation that show ways of meaningfully including stakeholders in the design, implementation, and interpretation of intervention research (HARC 2019; International Association for Public Participation 2017). These continuums encourage researchers to think about how community partners, including practitioners, families, and policy-makers, can be moved from a passive form of involvement (such as being fed information through infrequently assembled advisory committees) to active engagement in the conduct of the study.

Circling back to a point made at the beginning of this commentary, this will require ceding some level of control from the research side. Increased cooperation and communication can help researchers be sensitive to the burden that our designs put on stakeholders and how we attempt to mitigate these burdens. Researchers need to recognize the costs that come along with the benefits of studies designed to improve program practices and social policy (Frank et al. 2014).

Conclusion

All of this suggests the need for prevention researchers to be more strategic. Despite the recognized value and purpose of randomized trials and the oft-stated preference of funders and researchers for large-scale, multi-site controlled trials, there is increasing recognition of the challenges of translating the results of these types of studies to site-level work. This is partly not because of the pragmatic issues described here but also because averaging across sites can mask the variation seen between sites, reducing applicability of a global finding to the local setting (Orr et al. 2019). One result is that actual interventions that “work” in community contexts remain few and far between, and we cannot solve the problems of child maltreatment by being complacent with short lists of evidence-based prevention practices that have not been sufficiently field-tested. As Eddy and Sneddon (2019) note in the introduction to this section: “Science and practice will only move forward with the accumulation of reliable and valid knowledge about programs that are put into and shaped by practice”.

While we wait for new statistical methods to help us cope analytically with these challenges, we must also consider new research designs that embrace this complexity, work with our partners, and allow us to test, in faster fashion, innovative and tailored interventions that recognize the diversity of families within these programs and their different needs for help and support. Future work needs to focus on understanding the mediators and moderators of treatment effects, examining specific elements (or active ingredients) of child abuse and neglect prevention services that nudge the proximal outcomes and, in turn, lead to the more distal maltreatment reduction and child well-being outcomes emphasized by policy-makers and funders (see Supplee and Duggan 2019, for further discussion of active ingredients).

Increased interest in using continuous quality improvement methods such as the Institute for Healthcare Improvement’s Breakthrough Series Collaborative model (Kilo 1998) in early childhood home visiting (e.g., Arbour et al. 2019) is an example of how data-driven decision-making can be used by programs to test and enact service delivery adaptations in close to real time. But quality improvement is not, by itself, research. Rapid-cycle, iterative study methods such as SMARTs can merge the empirical rigor of experimental trials with the flexibility of service adaptation and may provide a path forward for future investigation. Note the use of the word “may” in the previous sentence. Although these emerging methods show promise, there are few examples specific to child abuse prevention programming.

As the four studies in this special section have demonstrated, we need to consider evaluation designs that go beyond our standard definition of effectiveness trials. These designs must address the nuances of service implementation in moving

proximal signifiers of change in order to help us better understand what really works for different kinds of families. Given the wide variety of services and support available in many of our communities for vulnerable families and the lack of consensus on which approaches are most helpful, almost nowhere is the need for this type of work clearer than in the field of child maltreatment prevention.

Funding Information The author's work was supported by the Health Resources and Services Administration (HRSA) of the US Department of Health and Human Services (HHS) under cooperative agreement UD5MC30792, Maternal, Infant, and Early Childhood Home Visiting Research and Development Platform. This information or content and conclusions are those of the author and should not be construed as the official position or policy of, nor should any endorsements be inferred by HRSA, HHS, or the US Government.

Compliance with Ethical Standards

Conflict of Interest The author declares that he has no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by the author.

Informed Consent Because this article is a commentary, informed consent is not applicable.

References

- Arbour, M., Mackrain, M., Fitzgerald, E., & Atwood, S. (2019). National quality improvement initiative in home visiting services improves breastfeeding initiation and duration. *Academic Pediatrics, 19*(2), 236–244. <https://doi.org/10.1016/j.acap.2018.11.005>.
- August, G., & Gewitz, A. (2019). Moving toward a precision-based, personalized framework for prevention science: Introduction to the special issue. *Prevention Science, 20*, 1–9. <https://doi.org/10.1007/s11121-018-0955-9>.
- Boller, K., Bradley, R., Cabrera, N., Raikes, H., Pan, B., Shears, J., & Roggman, L. (2006). The Early Head Start father studies: Design, data collection, and summary of father presence in the lives of infants and toddlers. *Parenting, 6*(2–3), 117–143. <https://doi.org/10.1080/15295192.2006.9681302>.
- Eddy, J. M., & Sneddon, D. (2019). Rigorous research on existing child maltreatment prevention programs: Introduction to the special section. *Prevention Science*.
- Eddy, J. M., Shortt, J. W., Martinez, C. R., Holmes, A., Wheeler, A., Gau, J., Seeley, J., & Grossman, J. (2019). Outcomes from a randomized controlled trial of the Relief Nursery program. *Prevention Science*. <https://doi.org/10.1007/s11121-019-00992-9>.
- Frank, L., Basch, E., & Selby, J. V., for the Patient-Centered Outcomes Research Institute (2014). The PCORI perspective on patient-centered outcomes research. *JAMA, 312*(15), 1513–1514. <https://doi.org/10.1001/jama.2014.11100>.
- Green, B. L., Ayoub, C., Bartlett, J. D., Von Ende, A., Furrer, C., Chazan-Cohen, R., Vallotton, C., & Klevens, J. (2014). The effect of Early Head Start on child welfare system involvement: A first look at longitudinal child maltreatment outcomes. *Children and Youth Services Review, 42*, 127–135. <https://doi.org/10.1016/j.childev.2014.03.044>.
- Green, B., Sanders, M. B., & Tarte, J. M. (2018). Effects of home visiting program implementation on preventive health care access and utilization: Results from a randomized trial of Healthy Families Oregon. *Prevention Science*. <https://doi.org/10.1007/s11121-018-0964-8>.
- Guastafarro, K., & Collins, L. M. (2019). Achieving the goals of translational science in public health intervention research: The multiphase optimization strategy (MOST). *American Journal of Public Health, 109*(S2), S128–S129. <https://doi.org/10.2105/AJPH.2018.304874>.
- Home Visiting Applied Research Collaborative. (2019). The importance of participatory approaches in precision home visiting research. HARC Research Brief [PDF file]. Retrieved from <https://www.hvresearch.org/precision-home-visiting/participatory-approaches/>. Accessed 15 Dec 2019.
- International Association for Public Participation. (2017). *IAP2 public participation spectrum* [PDF file]. Retrieved from <http://iap2usa.org/cvs>. Accessed 15 Dec 2019.
- Kaplan, D. (2019). *Bayesian inference for social policy research. OPRE report 2019: 36*. Washington DC: Office of Planning, Research and Evaluation.
- Kilo, C. M. (1998). A framework for collaborative improvement: Lessons from the Institute for Healthcare Improvement's Breakthrough Series. *Quality Management in Health Care, 6*(4), 1–13.
- Korfmacher, J., Green, B., Staerkel, F., Peterson, C., Cook, C., Roggman, L., Faldowski, R., & Schiffman, R. (2008). Parent involvement in early childhood home visiting. *Child and Youth Care Forum, 37*, 171–196. <https://doi.org/10.1007/s10566-008-9057-3>.
- LeCroy, C. W., & Lopez, D. (2018). A randomized controlled trial of healthy families: 6-month and 1-year follow-up. *Prevention Science*. <https://doi.org/10.1007/s11121-018-0931-4>.
- Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D., & Murphy, S. A. (2012). A "SMART" design for building individualized treatment sequences. *Annual Review of Clinical Psychology, 8*, 14.1–14.28. <https://doi.org/10.1146/annurev-clinpsy-032511-143152>.
- Marsh, J. C., Angell, B., Andrews, C. M., & Curry, A. (2012). Client-provider relationship and treatment outcome: A systematic review of substance abuse, child welfare, and mental health services research. *Journal of the Society for Social Work and Research, 3*(4), 233–267. <https://doi.org/10.5243/jsswr.2012.15>.
- Munns, A., Watts, R., Hegney, D., & Walker, R. (2016). Effectiveness and experiences of families and support workers participating in peer-led parenting support programs delivered as home visiting programs: A comprehensive systematic review. *JBIR Database of Systematic Reviews and Implementation Reports, 14*(10), 167–208. <https://doi.org/10.11124/JBIRIR-2016-003166>.
- O'Neill, K. M., Cluxton-Keller, F., Burrell, L., Crowne, S. S., & Duggan, A. (2018). Impact of a child abuse primary prevention strategy for new mothers. *Prevention Science*. <https://doi.org/10.1007/s11121-018-0925-2>.
- Ondersma, S. J., Martin, J., Fortson, B., Whitaker, D. J., Self-Brown, S., Beatty, J., Loree, A., Bard, D., & Chaffin, M. (2017). Technology to augment early home visitation for child maltreatment prevention: A pragmatic randomized trial. *Child Maltreatment, 22*, 334–343. <https://doi.org/10.1177/1077559517729890>.
- Orr, L. L., Olsen, R. B., Bell, S. H., Schmid, I., Shivji, A., & Stuart, E. A. (2019). Using the results from rigorous multisite evaluations to inform local policy decisions. *Journal of Policy Analysis and Management, 38*, 978–1003. <https://doi.org/10.1002/pam.22154>.
- Patsopoulos, N. A. (2011). A pragmatic view on pragmatic trials. *Dialogues in Clinical Neuroscience, 13*, 217–224.
- Paul, G. L. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology, 31*, 109–118.
- Peterson, C. A., Wall, S., Raikes, H. A., Kisker, E. E., Swanson, M. E., Jerald, J., Atwater, J. B., & Qiao, W. (2004). Early Head Start: Identifying and serving children with disabilities. *Topics in Early Childhood Special Education, 24*(2), 76–88. <https://doi.org/10.1177/02711214040240020301>.

- Sama-Miller, E., Akers, L., Mraz-Esposito, A., Zukiewicz, M., Avellar, S., Paulsell, D., & Del Grosso, P. (2018). *Home visiting evidence of effectiveness review: Executive summary*. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Washington, DC: Department of Health and Human Services.
- Schindler, H. S., McCoy, D. C., Fisher, P. A., & Shonkoff, J. P. (2019). A historical look at theories of change in early childhood education research. *Early Childhood Research Quarterly*, *48*, 146–154. <https://doi.org/10.1016/j.ecresq.2019.03.004>.
- Snow, R. E. (1991). Aptitude-treatment interaction as a framework for research on individual differences in psychotherapy. *Journal of Consulting and Clinical Psychology*, *59*(2), 205–216. <https://doi.org/10.1037/0022-006X.59.2.205>.
- Supplee, L. H., & Duggan, A. (2019). Innovative research methods to advance precision in home visiting for more efficient and effective programs. *Child Development Perspectives*, *13*, 173–179. <https://doi.org/10.1111/cdep.12334>.
- Supplee, L. H., Parekh, J., & Johnson, M. (2018). Principles of precision prevention science for improving recruitment and retention of participants. *Prevention Science*, *19*, 689–694. <https://doi.org/10.1007/s11121-018-0884-7>.
- Tosh, G., Soares-Weiser, K., & Adams, C. E. (2011). Pragmatic versus explanatory trials: The Pragmascope tool to help measure differences in protocols of mental health randomized controlled trials. *Dialogues in Clinical Neuroscience*, *13*, 209–215.
- Yousafzai, A. K., Aboud, F. E., Nores, M., & Kaur, R. (2018). Reporting guidelines for implementation research on nurturing care interventions designed to promote early childhood development. *Annals of the New York Academy of Science*, *1419*, 26–37. <https://doi.org/10.1111/nyas.13648>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.