



*Making the Legal System Work for
Children and Parents*

Introduction to Evaluating the Impact of Legal Representation Programs on Outcomes for Parents and Children

by Mark E. Courtney, PhD, Samuel Deutsch Professor
The University of Chicago, School of Social Service Administration

Background

There are strong legal principles supporting the position that parents and children involved with the child welfare court system in the United States should be provided with high quality legal counsel. In addition, many professionals in the child welfare field believe that high-quality legal representation for children and parents will lead to improved outcomes for families.

Unfortunately, legal principles often are not enough to drive public investment. Over the past two decades, the field of social science research has developed substantially, and new models of legal representation programs are expected to be able to prove their worth through industry-standard social science research. Conversely, the existing empirical evidence in support of the proposition that legal representation improves outcomes of interest to parents, children, and policymakers remains weak.

While there is growing consensus nationally regarding what constitutes “high-quality” legal representation for parents and children involved in the child welfare system, there is little consistency throughout the country in terms of practice. Therefore, efforts to improve and build the case for effective legal representation of these parties should include rigorous evaluation of existing and proposed programs.

Empirical evidence can provide a targeted basis for making effective improvements to existing and proposed programs. In addition, rigorous evaluations can engender government support and sustainability for them, as state and federal funding prioritizes evidence-based programs.

Purpose of this brief

This brief provides developers and operators of legal representation programs for parents and children involved in the dependency court system with advice regarding the conditions necessary for rigorously evaluating the impact of their programs on outcomes for the parents and children they serve. By outcomes, here we mean any change in the experience of parents and children

hoped to be affected by a legal representation intervention. These might be permanency outcomes or more immediate perceptions of system fairness. The principles described here apply equally to programs representing parents and those representing children.

- This is an introductory overview of key issues to consider when contemplating evaluation of the impact of your program, focusing exclusively on evaluation methods that have potential to establish *causal evidence of program impacts*. As programs are implemented and expanded over time, attending to these issues can enhance the ability of programs to establish evidence of their effectiveness.
- Programs seriously considering an impact evaluation are advised to seek the advice of an experienced evaluator well versed in the evaluation methods described here before embarking on an evaluation.

How can we evaluate the impact of legal representation on child and parent outcomes?

Even if they are convinced of the righteousness of their work, at some point all developers of legal representation programs ask themselves the following question: Does our program really improve outcomes for the parents and children affected by our program? Before we can design an evaluation that can answer that question, we need to clarify how a given program might affect outcomes for parents and children.

How might legal representation affect outcomes?

There are two primary mechanisms through which legal representation could influence outcomes for parents and children. First, and most obviously, representing these individuals could affect them by providing them with assistance that they would otherwise not receive. This would be a *direct* effect of legal representation. For example, the advocacy of an attorney who is able to engage with, assess the needs and desires of, and zealously represent a parent whose child has been placed in foster care may improve the likelihood that the parent receives services that assist the parent in regaining custody of their child.

Second, legal representation of individuals could, over time, influence the operation of the dependency court more generally, thereby improving outcomes for all individuals whose involvement with the child welfare system is overseen by the court regardless of whether they themselves are represented. This would be an *indirect* effect of legal representation.

This brief focuses on the most rigorous methods for establishing that legal representation *causes* changes in outcomes for parents and children.

Causation or association?

Program operators often use observations about the association between involvement in their program and outcomes experienced by program participants to argue that their program has had a positive impact on program participants. For example, an operator of a program representing parents involved in dependency proceedings might point to evidence that the rate of family reunification went up after the program went into effect to argue that the program caused the reunification rate to improve. Or the developer of a program that provides legal representation for

some children in foster care in a jurisdiction might compare the rate of placement instability between the children it serves to those it does not. Noting that the children served by the program experienced fewer moves than those it did not serve; the program developer might assert that the program caused a reduction in placement moves. In both cases, these program developers will have made flawed arguments on behalf of the impact of their program; *association is not the same thing as causation!* In the case of the parental representation program, the developer has inferred from the timing of the implementation of the program that an improvement in outcomes was caused by the program. However, there may have been many other developments happening during the same period (e.g., changes in other aspects of the child welfare or court systems; changes in the population of families coming to the attention of child welfare authorities; changes in the community context) that actually caused the change in outcomes. The timing of program implementation might be associated with the timing of an outcome, but that does not mean that the program caused the outcome. In the case of the children's legal representation program, the program developer assumes that the only meaningful difference between the children served by the program and those not served is the help provided by the program itself, but the process leading some children to be selected into and others to be excluded from the program could lead to many other differences between the two groups of children, differences that actually drive the observed differences between the groups in placement mobility.

Evaluation designs that help establish a causal relationship between legal representation and outcomes

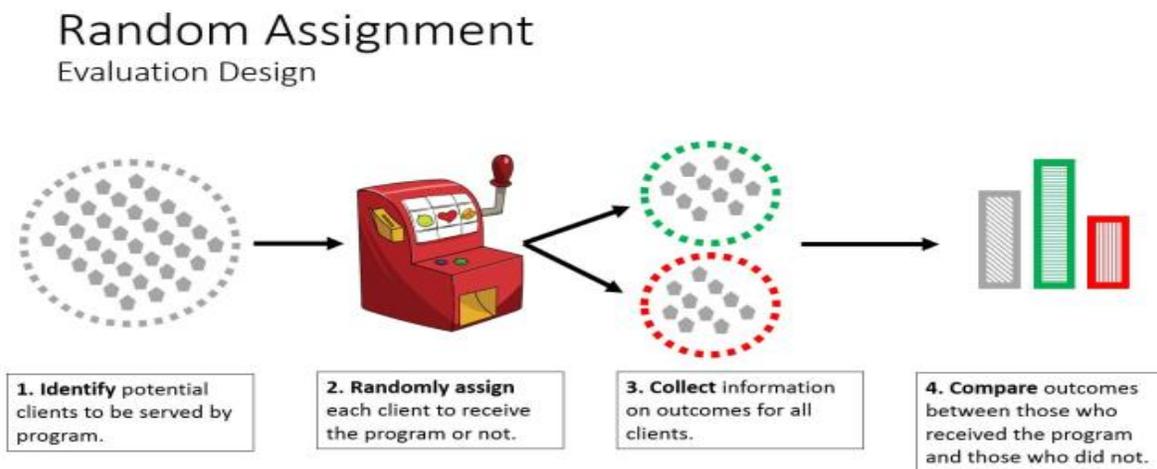
To build a strong case for the benefits of high-quality legal representation, program developers need to work closely with evaluators to design and implement program evaluations that can make a *causal* argument for the impact of legal representation programs on outcomes for parents and children. In other words, we want to be able to identify changes in outcomes that are directly attributable to a given program.

Given the nature of legal representation programs, there are two types of evaluation designs that have the greatest promise for strengthening the evidence base for legal representation. First, program developers should seek opportunities to evaluate their programs through random assignment of individuals to their programs while other individuals are randomly assigned to the existing or alternative forms of legal representation. Second, program developers should seek opportunities to implement their programs at the site level (e.g., county court systems; court calendars within a large county court system), ideally by randomly assigning some sites to provide the program and others to not provide the program.

Evaluating the impact of legal representation on individuals using random assignment

Random assignment evaluation designs use a process akin to a lottery to assign some members of a program's target population to be offered the program ("program group") and others to be excluded from participation in the program ("the comparison group"). Because the process of selection into receiving or not receiving the program is random, the characteristics of the two groups should differ only by chance alone. As a result, observed differences in later outcomes between the program group and comparison group can be attributed to the presence of the

program. Random-assignment evaluations allow the program developer to answer the following question: How do outcomes change for the target population of my program as a direct result of my program being offered to them? Note that the impact assessed using these evaluation designs is the impact on the target population (i.e., the entire population that is eligible for the program), not the population that is actually engaged by or completes the program. Thus, these are called *intent-to-treat* evaluation designs, since the population of interest is everyone the program would like to engage in services. Understanding this is key to deciding at what point in the process of involvement with the program random assignment should take place. For example, a program that is intended to provide legal representation to all indigent parents involved in dependency court proceedings will soon find that some of those parents choose not to accept the offer of representation. A random-assignment evaluation that assigned all indigent parents to the program or to a control group prior to program contact with parents would answer the following question: What is the impact of parental legal representation on outcomes for the population of indigent parents involved in dependency proceedings? In contrast, an evaluation design could call for first offering legal representation to all indigent parents involved in dependency proceedings and then randomly assign only those parents who accepted the initial offer of representation to the program or control group. A random-assignment evaluation that assigned indigent parents after they had accepted an offer of services would answer the following question: What is the impact of parental legal representation on the outcomes for indigent parents involved in dependency proceedings who indicate a willingness to be represented?



While many factors can come into play when considering a random-assignment program evaluation, several conditions are necessary for a feasible and rigorous design. First, there must be adequate excess demand for the program over a sufficient period of time to allow for assignment of program and comparison groups that are large enough to reliably assess program effects. The size necessary to assess program impacts depends largely on how large an effect will be considered substantively meaningful. The larger the number of individuals assigned to the program and comparison groups, the smaller the size of the effects that the evaluation will be able to reliably detect. The presence of excess demand for a program also addresses ethical concerns about denying what is perceived to be a valuable

service, in this case legal representation, to a population arguably in need of that service. If a program cannot serve the entire population of indigent parents that it would opt to represent, then it is likely to already be involved in rationing the provision of a scarce service in ways that are less equitable than random assignment. Second, the program's target population (i.e., the population to undergo random assignment) should be clearly defined and remain consistent over the course of the evaluation. Third, the program itself should be clearly defined and consistently delivered over the course of the evaluation. And fourth, the process of assigning individuals to the program and control groups must be clearly defined and adhered to over the course of the evaluation. Violations of group assignment (e.g., individuals who are assigned to the comparison group instead ending up receiving the service) are particularly problematic. An experienced program evaluator can help you assess whether your program is likely to be able to meet these conditions adequately enough to successfully evaluate your program using this approach to impact evaluation.

Evaluating the impact of legal representation through site-based implementation of a program

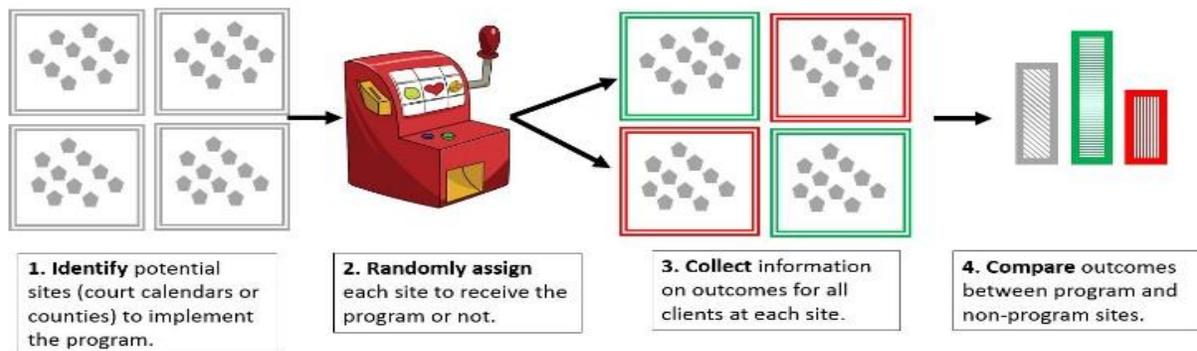
Adding a program of high-quality legal representation of children or indigent parents to dependency court proceedings could, over time, lead to changes in the behavior of all participants in those proceedings. Indeed, advocates expect that one mechanism through which quality representation will improve outcomes for parents and children involved with the child welfare system is through increasing the sensitivity of all members of the court team to the rights and needs of children and parents. If this assumption is correct, then implementing a program of legal representation has the potential to improve outcomes for all individuals who come into contact with a court in which the program operates, even individuals who for various reasons are not represented.

Therefore, to build a case for the benefits of high-quality legal representation programs, it is important to compare the outcomes of parents and children served by courts where these programs exist to the outcomes experienced by parents and children served by courts without such programs. However, in making these comparisons we will want to make sure that any differences we observe between outcomes in the "program" and "non-program" courts are a result of the program and not due to other differences between the courts or the courts' contexts. We need to be able to assure ourselves and others that what might appear to be effects of a program are not due instead to differences between courts that are relatively stable over time (e.g., differences in the size of the population served, such as between urban and rural court systems) or differences that evolve over time (e.g., the implementation of policy or practice reforms that are unrelated to the implementation of the legal representation program).

Evaluations that assess the impact of programs delivered to a population within a particular site (e.g., county or courtroom) are called *site-based evaluations*. These evaluations compare outcomes between sites that operate a program and those that do not, answering the question: What benefits accrue to parents and children as a consequence of being involved in a child welfare jurisdiction where high-quality legal representation is being routinely offered? Ideally, site-based evaluations are designed before a program is implemented in any site, with sites then being randomly assigned to implement the program ("program sites") or not ("non-program sites"). For example, a state that was implementing a program of legal representation for indigent parents but only had sufficient funding to provide the program in half of the counties in the state could randomly choose the program counties from among all counties in the state. Likewise, a large urban county with a dependency court that operated many

separate court calendars could choose to implement a program of legal representation for children in only half of the calendars, randomly choosing which would implement the program. Because the sites (counties or calendars) had been chosen randomly, after implementation of the program any observed differences between the program and comparison sites in outcomes for parents and children could be attributed to the presence of the program.

Site-Based Random Assignment Evaluation Design



In many cases it is not possible to randomly assign sites to implement a program. For example, a program developer may want to conduct a site-based evaluation, but decides to do so only after the program already operates in several counties that volunteered for the program when it was first proposed. Or perhaps the funders and developers of a program decided to start off small as they initially implemented their program, favoring small jurisdictions over larger ones. In either case, the choice of initial sites in which to implement the program might have favored sites whose outcomes already differed from those in sites where the program was not implemented; under these scenarios, comparing parent and child outcomes between the program and non-program sites would not make a very compelling argument for the impact of the program on outcomes.

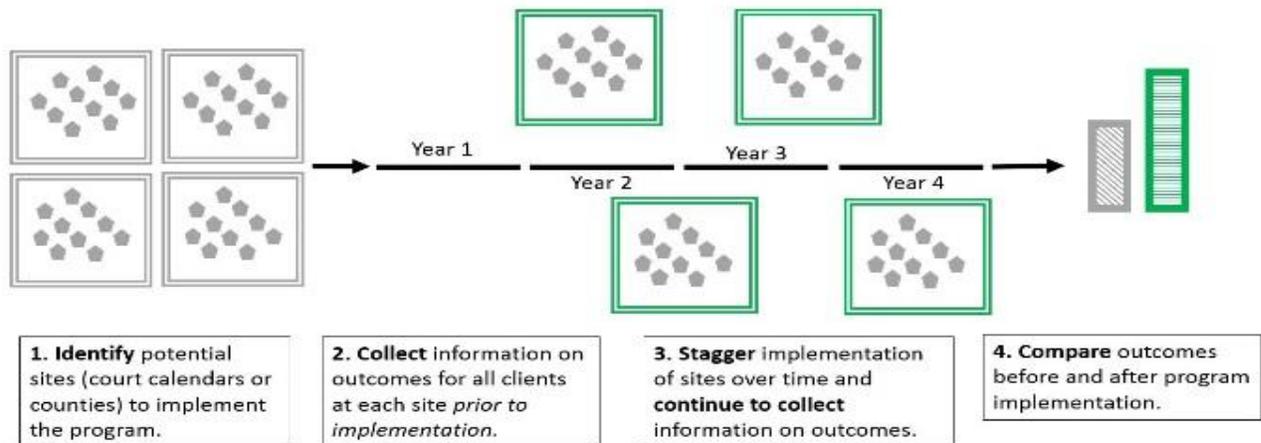
Nevertheless, it is possible under certain circumstances to rigorously assess differences in outcomes between program and non-program sites, without having randomly assigned the sites to program or non-program status, using what are referred to as *difference in differences* (DID) evaluation designs. DID evaluations take advantage of the staggered implementation over time of programs across multiple sites, comparing changes over time in outcomes between program and non-program sites. They look for changes in outcomes that are observed to occur at the point that a program is implemented within a given site that do not occur in non-program sites during the same time period. The assessment of the rigor of these evaluation designs relies heavily on the evidence that program implementation, rather than other changes simultaneously taking place in program sites, is driving observed changes in outcomes.

The factors that come into play when considering a site-based evaluation differ from those that apply to evaluation based on random-assignment of individuals. First, the ability of a site based evaluation to

reliably identify impacts of a program on child and parent outcomes depends heavily on both the number of program and non-program sites involved in the evaluation and the number of individuals involved in each site. In general, the larger the number of sites involved and the larger the number of individuals involved per site, the smaller the size of the effects that the evaluation will be able to reliably detect. Second, the program’s target population should be clearly defined and remain consistent over the course of the evaluation. Third, the program itself should be clearly defined and consistently delivered over the course of the evaluation.

Difference in Differences (DID) evaluation designs have some additional requirements that stem from the reliance of DID methods on assessing changes in outcomes over time, between and within sites, from before until after the implementation of a given program. First, they require that an evaluator has access to data on the outcomes of interest prior to and after program implementation.¹ For example, an evaluation might use child welfare administrative data on the timing of children’s exits to legal permanency across all counties in a state from before the program was implemented in any county until after the program had been fully implemented in all program counties. Second, they benefit from maximizing the staggering of program implementation over time between sites. For example, staggering the implementation of a program in ten counties within a state by implementing the program in two counties per year over a five-year period would make for a stronger evaluation than doing so in five counties in year two and five more in year four of a five-year evaluation. An experienced program evaluator can help you assess whether your program is likely to be able to meet the conditions for a DID evaluation adequately enough to successfully evaluate your program using a DID design.

Site-Based Difference in Differences Evaluation Design



¹ Site-based evaluations that randomly assign sites to program and non-program status do not require data on outcomes prior to program implementation, but the availability of such data can enhance such evaluations.