



ANNIE E CASEY FOUNDATION'S  
ADVOCACY LEARNING LAB

DATA ANALYSIS>  
DATA METHODOLOGY>  
DATA UNDERSTANDING

**Staff and Leadership understand the benefits, risks and limitations of using data to inform policy advocacy, including understanding the importance of identifying critical questions to explore using data.**

Cheryl McAfee and John Hodgins

# Why it Matters:

Data undoubtedly reveal a lot about child well-being, and the appropriate use of data resulted in important policy improvements for children time after time. But the use of data also carries risks. Advocates who understand those risks can minimize them (that is, avoid misinterpretations or hide problems).

# How to do it:

## Benefits of Data

Data reports generally have appeal across the political spectrum. Conservatives like the way that data can be used to provide accountability for government programs. Liberals appreciate the way that data illuminate problems they care about. Even when there is disagreement about the best solution to a policy problem, data can serve as a conversation starter, kicking off a policy debate that needs to happen.

Advocates need to be able to analyze data to inform their organization's policy agenda. There are at least six important ways that advocates regularly use data. Each of these has significant benefits.

- Monitor child well-being: They use data to study the well-being of children. This research can guide their decisions about what to include on their organizational agendas. And when advocates make data readily available, other researchers, journalists and policymakers can also use it to assess child well-being.
- Set goals: Advocates use data to set goals and benchmarks for child well-being and to assess whether states and other governmental entities are meeting those goals.
- Identify problems: They use data to investigate whether anecdotal stories are really symptomatic of a systemic problem, and what that systemic problem might be. This may inform their agenda and also assist in identifying solutions or policy improvements.
- Start a conversation: They use data to provide a compelling story that will help build public awareness, elevate the issue and draw policymaker attention; this is often the first step toward an advocacy success.
- Evaluate policy and programs: They use data to reveal whether particular approaches to the problem have been successful, so that they can determine whether more or different action is needed. Since many factors affect the well-being of children, it can be difficult to isolate the role of any particular policy in affecting well-being, but often the use of data is the only way to assess it, however limited that assessment may be.
- Rebut arguments: Advocates also use data and data analysis to rebut opponents' arguments or assertions in the list above.

# How to do it:

There are a number of approaches to analyzing data that can lead to important insights for policy advocacy. The kinds of comparative approaches listed below show some of the benefits of using data. They include:

- Comparisons over time can reveal whether a situation is getting better or worse. Comparisons across or within states can reveal that a state or community has a particularly bad (or good) situation, and can shed light on whether policies are in place that lead to better or worse outcomes.
- Comparisons across groups (e.g., by race and ethnicity, dual language learner status, disability and age group) can reveal disparities in outcomes due to policies and can help identify where resources should be targeted. For example, age group comparisons can show at what point children start falling behind educationally or start missing primary health care appointments.
- Comparisons by program or service can show that children who are known to be needy are accessing some services but not others.

# How to do it:

The Resources section includes case studies that show how the use of data directly led to improvements in child well-being. Data analysis provides more detailed information about how to analyze data and Data Visualization provides information on data presentation.

## Limitations of Data

As valuable as data can be, they have important limits. Advocates need to recognize and articulate these limits. **Statistical significance**

Sometimes, data may simply not be statistically meaningful because of small sample size. Advocates can sometimes group data over multiple years to expand the sample size and obtain more reliable information.

For example, the yearly numbers of infant deaths in many counties are not large enough to provide a reliable infant mortality rate. Combining several years of data provides a larger sample that produces more reliable results. Other times, with very large sample sizes, there may be statistically reliable results but there are simply no important differences between groups. For example, the difference between the poverty rate for males and females age 0 to 4 in the 2012 American Community Survey (49.2 percent, compared to 50.8 percent) is statistically valid because it is based on a large sample, but the difference is tiny and not important. A more challenging problem can arise when some data in a set are statistically meaningful and other data are not. While researchers sometimes provide all the data and flag the ones that are not statistically meaningful, the flags are often ignored.

Sometimes it is a better practice to just withhold data that are not statistically meaningful. However, there may be cases where showing that a difference is not statistically significant is important because people will assume the opposite without data. In other cases, as limited as the data may be, they may be the only data available and therefore the best available. In some cases, advocates may want to provide data that have important limitations because they are all that are available to inform policymakers about a situation. Advocates need to consider the expertise of the audience when making this decision, keeping in mind that many policymakers and influential people have very poor understanding of statistics. KIDS COUNT national publications do not mention statistical significance to avoid confusion among KIDS COUNT audiences, but the KIDS COUNT Data Center does include confidence interval files for most of the indicators from the American Community Survey (ACS). Thus, policymakers, journalists and others should not get confused, but researchers can find the information.

# How to do it:

## Obscure differences among groups

Data limitations can also obscure differences among groups of children, in ways that may distort policymaking.

- Data sets that fail to collect racial, ethnic, income or linguistic data, or data on disabilities may obscure significant differences in outcomes among groups of children. So may data where the sample size is too small to detect statistically significant differences between groups (or over time). If advocates use such data, it is incumbent on them to note that these differences may exist and to seek out other evidence where possible to provide a fuller picture. For example, if the state only has statewide data on lead poisoning, but some communities have much higher rates of hospital admission for lead treatment than others, that information would suggest that there is great variation among communities.
- Data that only measure the well-being of children receiving services (such as through health insurance or public education) tell us nothing about the children who are not receiving these services. For example, data on how many children in public insurance programs receive the recommended level of primary care visits do not reveal anything about uninsured children or children with private insurance.
- Data that only examine school-age children leave out children in their youngest and most vulnerable years.

Data that are collected infrequently, or collected in inconsistent ways that invalidate comparisons across time, are useless for assessing whether policy changes have affected the well-being of children.

# How to do it:

## Lack of understanding among audiences

Data have other limitations. There are a sizable number of people in any audience who will not understand data, or will not believe them. Sometimes this is because the data are too complicated. But increasingly, research is showing that the human brain is hardwired to reject evidence that contradicts something the person believes in strongly.

These studies do not mean that data are useless. Instead, they show us that data can be helpful when the person receiving the information finds it value-neutral, or believes that it supports their values. However, when the information directly conflicts with their values or their political beliefs they will reject it. For this reason, framing messages that start with a value shared by the target audience is critical. Message framing is discussed in **Effective Framing**. Some areas of child well-being are hard to measure.

Another limitation of data is that there are parts of child well-being that cannot be easily measured, at least using current techniques. For example, there are no widely accepted methods for regularly assessing young child mental health. Nor is there a good technique for assessing whether states are making the right decisions in child protective cases. While it is possible to determine how many children are placed in foster care, how many are left at home with services and how many children receive no intervention, it is not possible to determine whether the right decisions were made in individual cases. **Data alone cannot achieve policy change.**

Finally, while data are important, data and research are not sufficient to achieve policy gains. They must be partnered with strategic communications and advocacy.

The reason communication is such an important topic for data-based child advocates is captured by a statement from the Organization for Economic Co-operation and Development (Giovannini, 2008, page 14); "When good statistics exist, they too often go unnoticed or misunderstood by a broad audience." In other words, producing good data is not enough to ensure that the data are used or used properly. *Storytelling, Communicating Data and Framing Disparities Data* provide guidance on how to communicate data effectively.

# How to do it:

## Risk of Using Data

Data collection and analysis can require a lot of staff time and tools. Relying too much on data can be costly and may not always be the best way to maximize the organization's resources.

With limited resources, it is important to balance the resources that go toward data collection and analysis with the resources that go toward dissemination and presentation. Too often, the resources dedicated to a study are consumed by data collection and analysis leaving too little for getting the information out to users. It is important to assess this balance at the beginning of a project rather than waiting until the data analysis is completed.

When advocates ignore the limitations of data, or fail to explain them and put them in context, the result can be damaging. This may be particularly true when reporting disparities in race and ethnicity without context.

For example, reporting the higher incarceration rates of minority youth, especially boys, without also explaining how at every decision point the system is structurally designed to treat minority youth worse, can actually increase the negative stereotypes about minority youth that they are violent, untrustworthy and dangerous.

If advocates report data inaccurately, or fail to note the limits of the data, that can put the organization's credibility at risk.



# How to do it:

## Forming a Research Question

One of the reasons that understanding the benefits, risks and limitations of using data matters is because it is important to take these into account when developing a good research question. The decision to investigate a topic might be prompted by:

- a change in federal law, which opens the door to new state action.
- a review of data that shows something unusual or worrisome in a trend or pattern (e.g., one Alabama research project began when they realized that their state consistently had the highest child death rate in the nation).
- the existence of "outliers" that might reveal important differences in policy or practice.
- current state policy debates. It might be designed to investigate whether a case exists to make an argument for or against a particular policy position.

Whatever prompts the research, the research questions must be carefully developed in order to get meaningful results.