Racial disproportionality in child welfare: False logic and dangerous misunderstandings

Jesse Russell
Racial disproportionality in child welfare: False logic and dangerous misunderstandings

Abstract: Disproportionality and disparities in child welfare appear to be widely recognized, if not fully understood, phenomena. There is often disagreement on how to interpret or find meaning in the empirical evidence that supports the existence of disproportionality and disparities—some the result of fertile and valuable discussion, some stemming from misunderstanding. Several potential paths of misinterpretation are examined here: the ecological fallacy concept, the fallacy of hidden assumptions, the lessons from different measures of disproportionality, the difficulty in understanding how probabilities relate to each other, and the effect that multicolinearity can have on statistical findings. Ultimately, better understanding of empirical findings helps to develop with confidence the tools, strategies, and initiatives for reducing disproportionality and disparities.
Disproportionality and disparities in child welfare appear to be widely recognized, if not fully understood, phenomena.\textsuperscript{1} Dependency court judges see first-hand racial imbalances appearing in the courtroom and struggle to understand why and how to safely reduce the imbalance to achieve the best outcomes for children and families. There is often disagreement on how to interpret or find meaning in the empirical evidence that supports the existence of disproportionality and disparities; some the result of fertile and valuable discussion; some possibly stemming from honest misunderstanding or misinterpretation of the empirical evidence. Several of these potential missteps are addressed in this paper, and these might not be the only potential pitfalls in meaningfully understanding the data surrounding disproportionality and disparities. They are presented here merely to further the discourse along a productive path.

Clarifying areas of misunderstanding or misinterpretation of data is a step toward formulating policies, practices and tools for reducing disproportionality and disparities in a meaningful way. This paper introduces five areas of potential misinterpretation or misunderstanding of empirical data, and it certainly does not exhaust these topics. The five areas of potential misunderstanding or misinterpretation are based in:

- the ecological fallacy concept,
- the fallacy of hidden assumptions,
- the lessons from different measures of disproportionality,
- the difficulty in understanding how probabilities relate to each other, and
- the effect that multicolinearity can have upon statistical findings.

This paper fits into a larger discussion about evaluating research, policy and best practices on racial disproportionality and disparities in juvenile dependency. First, some definitions are in order. For this paper, disproportionality is defined as follows: disproportionality refers to one population that is out of proportion with respect to the general population. Disparity is defined as a lack of equality: unequal treatment of one racial or ethnic group as compared to another racial or ethnic group. The Courts Catalyzing Change (CCC) Steering Committee expressed the judicial perspective on the challenges of communicating understanding about disproportionality and disparity as:

The challenge is how to maximize this opportunity to do something that will reduce disparity without getting mired down in the feelings and emotions you have when you think about how this affects one personally. The challenge is to stay focused and the question is: How do I communicate this in my jurisdiction in a way that will not create barriers.

The judges are right—the challenge is to maximize the opportunity. The challenge is not only to do so without being stuck in the emotional mire, but also without being stuck in any empirical/statistical swamps.

---

3 Id.
4 The Courts Catalyzing Change: Achieving Equity and Fairness in Foster Care Initiative has been organized by the Permanency Planning for Children Department of the National Council of Juvenile and Family Court Judges in partnership with Casey Family Programs, and is supported by the U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
1. The Ecological Fallacy

It has been posited that the large representation of African American children in foster care is due to higher maltreatment rates for African American children, resulting in the disproportionate representation of African American children in the dependency system. This claim must be carefully examined to fully understand what the actual units of analysis are, what are the groups being analyzed, and who are the individuals making up those groups.

For example, Elizabeth Bartholet argues that: “First and foremost is that blacks [sic] are disproportionately associated with a set of characteristics that have been repeatedly found by many different child welfare experts to be accurate predictors for child maltreatment … and there is no doubt that they are disproportionately associated with black families.”\(^6\) She argues that “there is substantial evidence that black maltreatment rates are significantly higher than white, because black families are affected by poverty and other risk factors for maltreatment at significantly higher rates than whites.”\(^7\) At first blush, this logic seems obvious: some characteristics are associated with higher rates of maltreatment, one group is more likely to have these characteristics, and thus, this group is more likely to have higher rates of maltreatment. However, underlying this argument is a logical fallacy. The fallacy does not mean that the conclusion is necessarily wrong, but it does mean that the conclusion does not actually follow from this argument. This error is known as the ecological fallacy.

The ecological fallacy occurs when one assumes that what holds true for a group also holds true for individuals within that group. This is based on the idea that a correlation (or statistical relationship) between two factors that describe group averages are ecological correlations. These ecological correlations can be different from the correlations that occur\(^6\) Elizabeth Bartholet, *The racial Disproportionality Movement in Child Welfare: False Facts and Dangerous Directions*, 51 Arizona Law Rev. 871 (2009).\(^7\) Id. at 900.
among the individuals within the group. An ecological fallacy occurs when someone mixes up an ecological correlation for a correlation among individual members of a group.

For example, often in presidential elections, states that are wealthier on average tend to vote Democratic. In those same elections, however, wealthier individual voters tend to vote Republican. If we looked only at the ecological correlation (wealthier states tend to vote for Democratic candidates as a state) and assumed it was true for individuals as well, we would be wrong. It is an ecological fallacy to assume that the ecological correlation that exists broadly also applies among individual cases more narrowly. The ecological fallacy does not mean that the two correlations (the aggregate/ecological and the individual) necessarily oppose each other. However, it does mean that you cannot draw conclusions from the group averages to apply to the individual members.

Robinson analyzed census data on literacy rates and the percentage of the population born outside of the United States.\(^8\) He showed that if you look at the average rates of literacy by state and look at the average proportion (the group) of the population that was foreign-born in each state, then there would be a strong positive correlation (0.53) between states with more foreign-born residents and states with higher literacy rates. The ecological fallacy would lead one to conclude from this that the members of the population who are foreign-born were more literate than the rest of the population. In fact, Robinson’s data showed that if you looked instead at the relationship between being foreign-born and illiteracy among individuals, foreign-born residents were actually less likely to be literate (the correlation was -0.12). Why did these two correlations point to opposite relationships? The reason that the relationship at the individual level was different from the relationship at the group level was that foreign-born individuals

---

were more likely to move to areas where the literacy rate was higher and these new residents did not themselves increase the local literacy rates. The point of this example is to show how the wrong conclusion can be reached if correlations are confused with each other.

With respect to disproportionality and disparity, Bartholet’s threshold argument is that the average risk level relates to the average rate of maltreatment. This is likely true and correct for the \textit{group average} of the whole population. Put another way, this is an ecological correlation. Based on this type of statement, however, we do not know if it is also true for individuals or all individual racial groups who make up the population. The correlations she offers about the whole population (a “powerful connection repeatedly demonstrated between poverty and related risk factors and maltreatment”\textsuperscript{9}) does not in fact tell us what the correlation is like for the different racial groups who make up that population. Her argument relies upon a correlation for the \textit{whole population}, but her conclusion is about a \textit{member} of the population.

In sum, knowing that higher rates of risk characteristics tend to correspond to a higher incidence of maltreatment in the population, and knowing that one group has a higher level of risk characteristics, cannot be used to conclude that the group is more likely to have a higher level of maltreatment. We cannot conclude that African American families have higher maltreatment rates because they have higher risk characteristics, just as we cannot conclude that a wealthier person voted for a Democrat because they live in a wealthier state, and we cannot conclude that immigrant groups are more literate because they live in high literacy rate counties.

Evidence from the Fourth National Incidence Study\textsuperscript{10} may provide some important information to illuminate the discussion. The first three iterations of the National Incidence Study found no relationship between racial groups and incidence of neglect or maltreatment. The

\textsuperscript{9}Bartholet, \textit{supra} at 902.
fourth iteration of the Study did find higher rates of maltreatment in African American families than others.\textsuperscript{11} However, the study investigators point out that while the study does see racial differences, but they cannot tell if the differences they see by race are due to race.\textsuperscript{12} Thus, the researchers emphasize that one cannot conclude that families from one racial group are more than families from another racial group to be more likely to maltreat children.

2. Hidden Assumptions

Since the existence of disproportionality and disparities is an accepted fact, some discussion has revolved around how to understand what disproportionality and disparities actually represent: disproportionate needs, higher rates of risk factors, family characteristics, community and neighborhood traits, implicit bias in decision-makers, or structural racism.\textsuperscript{13} Most importantly, discussion includes the question of whether disproportionality and disparities mean too many African American children are spending too much time in foster care, too few White children are spending too little time in foster care, or is one group being served better than the other is.

Potential misunderstanding can come into play when examining disproportionality as either over- or under-representation. The existence of disproportionality alone cannot tell us if it is due to one group’s over-representation or due to another group’s under-represented. In fact, it is very difficult to tell from the types of data currently being collected what the correct level of representation in foster care might be. If disproportionality is defined as over- or under-representation, then there is an unstated normative assumption about the proper level of representation.

\textsuperscript{12} *Id.*
\textsuperscript{13} Hill, *supra* at 8.
representation and an implicit assumption that one level of representation is good, while another is not as good.

For example, Bartholet argues that, “if white children are not being removed to foster care at rates equivalent to black rates given the incidence of actual maltreatment, it means that white children are being disproportionately denied protection.”\textsuperscript{14} This is a problematic conclusion because there is a hidden assumption in the argument from which this conclusion is derived. The assumption is that better protection comes from more and longer stays in out-of-home care, with no evidence presented to suggest that this might be true. This is an empirical question, and so far, the evidence and research does not support the idea that more and longer stays in out-of-home care necessary lead to better outcomes.\textsuperscript{15} Without greater empirical clarification of the causes of disproportionality, there is no way to assess whether one group is getting “too much” protection or “too little,” or is in the system too much or too little.

The “hidden assumptions” error can also take the following form: since there is racial disproportionality in the out-of-home care system, then “rates of reporting, substantiation and removal of black children who are suspected victims” of abuse and neglect should be reduced.\textsuperscript{16} Again, this line of reasoning is faulted by its unstated and unverified assumptions about what outcomes should be encouraged, or what rates of reporting and removal there ought to be.

Without rigorous empirical tests of hypotheses about the proper rates of reporting, substantiation and removal using solid data, there is no basis from which to make these sorts of

\textsuperscript{14} Bartholet, \textit{supra} at 921.
\textsuperscript{16} Bartholet, \textit{supra} at 911.
conclusions—the assumptions incorporated in the conclusions are unsupported thereby undermining the conclusions themselves.

3. Measuring Disproportionality and Disparities

The third area of possible misunderstanding centers upon how levels of disproportionality and disparities are measured and compared. Disproportionality can be demonstrated by showing that the proportion of one racial group within care is greater (or less) than the proportion of that racial group in the community—for example, if the population of African American children in a county is 25,000 and the number of African American children in care in that county is 75, then the rate per 1,000 is 3. And if the population of non-African American children in the county is 475,000 and the number of non-African American children in care in that county is 475, then the rate per 1,000 is 1. Thus, you could say that African American children were three times more likely to be in out-of-home care than non-African American children.

This works well for revealing disproportionality and is the basis for the racial equity scorecard developed by the Casey-CSSP alliance for Racial Equity in Child Welfare.\textsuperscript{17} To analyze variances in disproportionality (necessary for understanding which programs, tools and strategies might reduce disproportionality), comparisons across sites and across time are also needed.

To understand disproportionality and to be able to form potential solutions to the problems of disproportionality, the causal factors related to disproportionality must be understood. To do this, accurate, reliable and valid measurement of not just the fact of

\textsuperscript{17} See Derezotes, et al. Evaluating Multi-Systemic Efforts to Impact Disproportionality through Key Decision Points (2008).
disproportionality, but also the variances in disproportionality across units, such as courts, jurisdictions, counties, states, years, or decision points is needed.

Shaw, et al., address this issue directly, explaining that, “[a]lthough increasing attention is being paid to the disproportional representation of children of color in the child welfare system, the question of how to best measure over and underrepresentation over time and across localities has not yet been resolved.”\(^{18}\) The point they make is a good one: how to create a measure of disproportionality that can travel as a concept from one place to another, from one time period to another, or from one decision process to another. So just because two courts show disproportionality scores of 3.0, does not mean that each court has the same causal factors leading to that disparity.

Shaw, et al. suggest another way of calculating disproportionality, which they label the Disparity Index.\(^{19}\) Comparing rates per 1000 or using the Shaw, et al., Index is a way of measuring one aspect of disproportionality, but which might not be the only aspect that we want to look into. This type of measure works especially well for telling us how much more likely one group is to be represented in out-of-home care compared to another group for a given locality. We can say that in this town African American children are twice as likely to be in out-of-home care as Asian children, or that in this state Latino children are one and a half times more likely to be in care than all other children.

To focus on broader causal explanations of disproportionality, we could investigate the variations across localities, across states, counties, cities and years. If we want to know how to address any systemic causes of disproportionality, we would need the variations in disproportionality over time in some places, or why two places, or institutions, or jurisdictions

---


\(^{19}\) Shaw, *et al.* offer an Excel spreadsheet for calculating disparity measures from underlying data. Available at http://cssr.berkeley.edu/ucb_childwelfare/DisparityIndices.aspx
have had different experiences with disproportionality, using a measure that describes how
disproportionate a system is in total, not only how much one group is represented compared to
another.

Different places have different racial groups which are represented within local foster
care systems and dependency courts at different levels. If we limit the concept of
disproportionality to only one racial group, or to only one group at a time, we might miss the
bigger picture. For example, one county in the Midwest might have relatively few African
American children in its juvenile dependency system and have relatively more Native American
children in care. Another county in California might have higher numbers of African American
children and higher numbers of Latino children. If we used only the rates per 1000 or the Shaw,
et al., Index, we would have no good way to compare disproportionality across these two
counties.

One possibility way to address this issue would be the common “coefficient of variation”
measure, which is a useful form of the standard deviation (Greene 2008). The coefficient of
variation provides a means of comparing information across cases or across time. It has a value
of zero when there is no disproportionality, which is intuitively helpful, and the number would
increase with greater disproportionality. The coefficient of variation would be first calculated as
the standard deviation of the rate per 1000 of children in care for each group within a locality,
and then that standard deviation would be divided by the average rate for that locality. This type
of measure could potentially help answer some important questions about broad trends in
disproportionality across many different locations and times.

21 The coefficient of variation is the standard deviation of the sample divided by the mean of the sample. The
standard deviation represents how much all the observations differ from the average of the observations. The
coefficient of variance abstracts from this so that these differences can be compared across different samples.
A related interpretation issue rests in the difference between measuring disproportionality in representation and measuring disparate treatment. There may be some usage differences with these terms, but the common idea is that “disproportionality refers to the difference in the percentage of a group of children in the child welfare system as compared to that group’s percentage in the general population,” while “disparity means that one group of children experiences inequitable treatment or outcomes as compared to another group of children.”\(^\text{22}\) As one pair of analysts points out, “disproportionality of children in foster care is a function of disparity in the entry and/or exit process.”\(^\text{23}\)

Disparities in child welfare play out through racial differences in access to appropriate and effective services (for children and parents) to more quickly reunify, and to minimize the length of out-of-home placement and the number of moves in out-of-home care. The challenge with establishing broad-based measures of disparities that can be applied across multiple locations, times, and decision points, is that the existence of and quality of data systems varies greatly. Some jurisdictions are able track a wide variety of key measures while some jurisdictions would be required to track measures by hand. To compound the difficulty, different terms and concepts are defined can mean different things in different jurisdictions, and are even defined differently across systems. As Giovani Sartori notes, “the wider the world under investigation, the more we need conceptual tools that are able to travel.”\(^\text{24}\) For example, pre-hearing conferences in one jurisdiction might focus on discovery, while in another jurisdiction pre-hearing conferences are about group decision making for the child or family. Hence “[w]e do


\(^{23}\) Wulczyn & Lery, supra at 5.

\(^{24}\) Giovani Sartori, Concept Misinformation in Comparative Politics, 64 The American political Science Review, 1033, 1034 (1970).
need, ultimately, ‘universal’ categories—concepts which are applicable to any time and place,” and these need to be empirically precise as to what we are measuring and comparing.

4. Probabilities

Probabilities can be misleading. For example, if thirty percent of all children in care are African American and seventy percent are White, does this indicate that African American children are more or less likely to be placed into foster care? The likelihood of being placed into care is conditioned by the likelihood of having a case brought before a juvenile dependency court. How do these probabilities relate? What can we say about the probability of a child being in foster care given that she is African American, compared to the probability that a child is African American given that she is in foster care? These are not the same probabilities and they have a very specific relationship to each other.

One relationship of probabilities is worth pointing out, often referred to as Bayes’ Theorem. The idea of the Theorem is that the probability of event A (e.g., rain tomorrow) given event B (e.g., the weatherman forecasting rain) depends not only on the relationship between A and B (i.e., the accuracy of the forecast) but on the absolute probability of A independent of B (i.e., how much rain the area gets normally). Estimating the likelihood of it raining tomorrow depends on the news forecast, the average accuracy of the news forecasts, and on how much rain the area gets.

---

25 Id. at 1035.
26 These probabilities do not take into account real-world time sequences, such as when a removal occurred, when a child was placed in care, and when a case came to court. Also, it is only in an abstract sense that we can talk about the probability of a child being African American. In the real world, of course, someone’s race is not a matter of probabilities or likelihood functions.
27 Bayes’ Theorem or Bayes’ Rule is named after Thomas Bayes who first explained the idea. See, inter alia, Andrew I. Dale, A History of Inverse Probability (1999); or John A. Hartigan, Bayes Theory (1983).
As a more topical example, consider statistics from the 2005 National Survey of Child and Adolescent Well-being which indicates that of children who are introduced to child welfare services, 32% of cases are substantiated and 27% of children are placed in out-of-home care\textsuperscript{28}. What is the relationship between these statistics? The report shows that 33% of cases involving African American children are substantiated and 33% of cases involving White children are substantiated. The report further indicates that of African American children in out-of-home care (including foster care, kinship foster care and group care), 69% of cases were substantiated. For White children in out-of-home care, 54% of cases were substantiated. Finally, 40% of African American children are placed in out-of-home care while 37% of White children are placed in out-of-home care.

Using these percentages as probabilities, we can see if for White children and for African American children the probabilities of being placed in care given that their cases were substantiated are the same. That is, is a White child with a substantiated case just as likely to be placed in care as an African American child with a substantiated case?

Using Bayes’ Theorem, we can calculate the probability that a White child will be placed in out-of-home care if the case is substantiated to be 0.61.\textsuperscript{29} We can calculate the probability that an African American child will be placed in out-of-home care if the case is substantiated to be .84. This means that African American children determined to be victims of child abuse or neglect are 38% more likely to be removed and placed in out-of-home care. Or put another way, for every 10 White children who are determined to be victims of child abuse or neglect, six will be removed from their homes and placed in care. But for every 10 African American children


\textsuperscript{29} Probability of being in care given substantiation = probability of substantiation given being in care times the probability of being in care, divided by the probability of substantiation = (.69 * .40)/.33 = .61.
who are determined to be victims of child abuse or neglect, eight will be removed from their homes and placed in care.

To be able to combine a long series of statistics into a single intuitive sentence—among substantiated cases, African American children are much more likely to be placed out of the home—can be helpful in moving the discussion forward.

5. Multicolinearity

Multicolinearity is a term that describes those instances when two factors are tightly related in the real world or when two variables are closely correlated.\textsuperscript{30} Discussion about the causes of disproportionality and disparities often relies upon the statistical tool of regression,\textsuperscript{31} but one of the rules of regression is that there is no multicolinearity among the explanatory variables.

Say we wanted to test the influence of eating in restaurants and the influence of eating meals high in salt on a person’s blood pressure. We might hypothesize that people who eat in restaurants frequently and people who eat saltier meals could have higher blood pressure. Regression would be a proper test to see what influence these factors have, except that restaurant meals are saltier meals. Restaurants typically add more salt to their dishes than people do at home. So the two things we are measuring (restaurant meals and salty meals) have a lot of overlap. If we include both of these factors in our regression, then we might not see any real effect for either of them. Since the regression does not take into account that the measures are closely related to each other, any effect that each might have on the outcome variable (e.g. blood

\textsuperscript{30} Greene, supra. 2008
\textsuperscript{31} The most common and most versatile estimator of linear relationships among statistical variables is the ordinary least squares regression. It is the most efficient un-biased estimator of linear relationships, but it does have some restrictive assumptions (such as non-multicolinearity, as well as normally distributed error terms, linearity of the relationships and constant variances).
pressure) is lost. Like two children trying to yell over the top of the other, it becomes impossible to hear either one.\footnote{In more technical terms, this is a problem of instability of the inverse matrix or the impossibility of computing a regular inverse the matrix.} The solution is to run the regression with one of the measures, not both.

Within the discussion of disproportionality and disparities, this becomes a problem when testing the effects of race and of poverty on disproportionality and disparities. Our measures of race and our measures of poverty correlate with each other, thus creating multicollinearity. This multicollinearity confounds the regression, and the effects of these two variables either swamp each other entirely, or the voice of one gets lost behind the louder voice of the other. The historical context and the institutional and social legacies of race make measures of race and poverty too closely related for both to be included as measures in a regression. This does not mean that they could not be both separately assessed, but it does mean that a lot of theoretical and statistical care is needed to do so.

Conclusions

The data tells us that there is something out of balance. Judges’ observations in their courtrooms reflect the imbalance. Understanding disproportionality and disparities in juvenile dependency along with examining the data, the measurement techniques, and the disagreements, helps ensure that efforts are made to find the best way to understand these questions. Ultimately, this means that we can be more likely to find meaningful and useful answers. It is important to examine carefully and meticulously the data, the conclusions drawn from the data, and the assumptions incorporated into the conclusions. Only then can we confidently develop the tools, strategies, programs and initiatives to reduce disproportionality and disparities, and best serve the children and families in the dependency system.