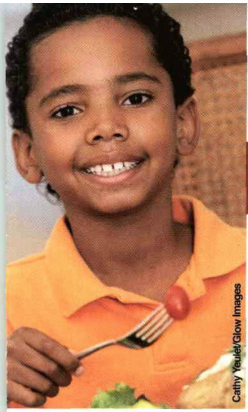
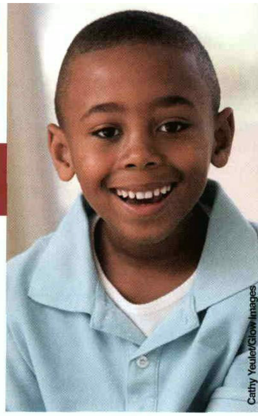
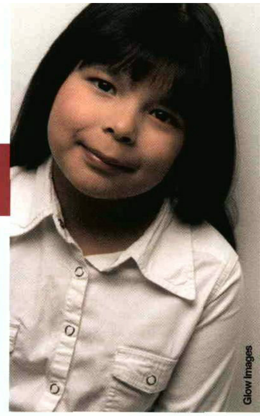
RACIAL DISPROPORTIONALITY IN CHILD WELFARE: False logic and dangerous misundersta
Russell, Jesse
Judicature; Nov/Dec 2011; 95, 3; ProQuest Criminal Justice
pg. 109







RACIAL DISPROPORTIONALITY IN CHILD WELFARE:

False logic and dangerous misunderstandings

by **JESSE RUSSELL**

Disproportionality and disparities in child welfare appear to be widely recognized, if not fully understood, phenomena.[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.

Dependency court judges struggle to achieve the best outcomes for children and families.

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 multicollinearity 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 being out of proportion with respect to the general population.[2] Disparity is defined as a lack of equality:[3] unequal treatment of one racial or ethnic group as compared to another racial or ethnic group. The National Council of Juvenile and

1. *Cf.* Dennette Derezotes, *et al. Race Matters in Child Welfare: The Overrepresentation of African American Children in the System* (2005); Government Accounting Office, *African American Children in Foster Care: Additional HHS Assistance Needed to Help States Reduce the Proportion in Care* (2007); Robert B. Hill, *Synthesis of Research on Disproportionality in Child Welfare: An update* (2006); Fred Wulczyn, *Racial Disparity in Foster Care Admissions* (2007).
2. Fred Wulczyn & Bridgette Lery, *Racial Disparity in Foster Care Admissions*, 5 (2007).
3. *Id.*

Family Court Judges' Courts Catalyzing Change (CCC)[4] Steering Committee expressed the judicial perspective on the challenges of communicating an 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 on the question: How do I communicate this in my jurisdiction in a way that will not create barriers?[5]

The judges of the CCC Steering Committee 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.

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 understand fully the actual units of analysis, the groups being analyzed, and 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 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 an analysis only considered the ecological correlation (wealthier states tend to vote for Democratic candidates as a state) and assumed it was true for individuals as well, it 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 conclusions cannot *be drawn 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 an examination of the average rates of literacy by state and the average proportion (the group) of the population *that was* foreign-born in each state would suggest 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 one 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 relationship at the individual level was different from the relationship at the group level, that foreign-born individuals were more likely to move to areas where the literacy rate was higher but these new residents did not themselves increase the local literacy rates. The point of this example is to illuminate 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 *group average* of the whole population. Put another way, this is an ecological correlation. Based on this type of statement, however, it is not known 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

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.

5. Sophia Gatowsky, *et al. Courts Catalyzing Change: Achieving Equity and Fairness in Foster Care—Transforming Examination into Action.* Juvenile and Family Justice Today 16, 19 (2008).

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.

8. William S. Robinson, *Ecological Correlations and the Behavior of Individuals,* 15 AMERICAN SCIENCE REV. 351 (1950).

risk factors and maltreatment"[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 *whole population*, but her conclusion is about a *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. It would be a mistake of logic to conclude that African American families have higher maltreatment rates because they have higher risk characteristics, just as it would be a mistake to conclude that a wealthier person voted for a Democrat because they live in a wealthier state, and it is a mistake to conclude that immigrant groups are more literate because they live in high literacy rate counties.

The ecological fallacy might have arisen in this situation based on an assumption that the population correlation between risk factors and maltreatment would apply equally to all subgroups within the population. It is offered as an empirical finding that risk factors correlate with maltreatment. However, it has not been established that these factors imply a similar risk across all groups within the population. It could be that one factor likely to increase the risk of maltreatment for White families is involvement with methamphetamines, while this is not as large a risk for African American families. Reciprocally, it could be that neighborhood characteristics play a larger role as a risk factor for African American families than for White families. (In fact, research has shown that neighborhood differences can have a disparate impact on families of color in terms of service availability and access.[10]) If the factors that increase risk of maltreatment are not identical across all racial groups, then it is a mistake to assume that broad correlations over the population would apply equally to all sub-sets of that population.

POTENTIAL MISUNDERSTANDING CAN COME INTO PLAY WHEN EXAMINING DISPROPORTIONALITY AS EITHER OVER- OR UNDER-REPRESENTATION.

Evidence from the Fourth National Incidence Study[11] relates to this discussion. The first three iterations of the National Incidence Study found no robust relationship between racial groups and incidence of neglect or maltreatment. The fourth iteration of the Study did find higher rates of maltreatment in African American families than in others.[12] However, the study investigators point out that while the study does see racial differences, they cannot tell if the differences they see by race are *due* to race.[13] That is, maltreatment rate differences that can be partly accounted for by dividing the sample into racial groups can be better accounted for by dividing the sample into economic groups instead. Economically disadvantaged families are more likely to experience maltreatment than economically advantaged families. African American families might be more likely to experience maltreatment, but primarily because they often belong to the economically disadvantaged group.

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.[14] 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 whether one group is being served better than the other is.

Potential misunderstanding can come in to 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*-representation. 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 and an implicit assumption that one level of representation is good, while another is not as good.

9. Bartholet, *supra* at 902.

10. D. E. Roberts, *Toward a community-based approach to racial disproportionality.* 22 PROTECTING CHILDREN 4 (2007); Fred Wulczyn, *Permanency, disparity and social context.* Presentation for the Race and Child Welfare Working Conference at Harvard Law School. January 28-29, 2011. Retrieved from http://www.law.harvard.edu/programs/about/cap/cap-conferences/rd-conference/rd.conference.papers.html

11. Andrea J. Sedlak, *et al. Fourth National Incidence Study of Child Abuse and Neglect (NIS-4): Report to Congress, Executive Summary* (2010).

12. Andrea Sedlak, *et al. Supplementary Analysis of Race Differences in Child Maltreatment Rates in the NIS-4* (2010).

13. *Id.*

14. Hill, *supra* at 8.

15. Bartholet, *supra* at 921.

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.*"[15] 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 leads to better outcomes.[16] Without this clarification, there is no way to assess whether one group is in the system too much or too little. Disproportionality has acted as an alarm that something in the trends is amiss, but it alone does not indicate that the problem is one of "too much" protection 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.[17] 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 conclusions—the assumptions incorporated in the conclusions are unsupported thereby undermining the conclusions themselves.

Measuring Disproportionality and Disparities

A third area of possible misunderstanding centers upon the ways in which levels of disproportionality and disparities are measured and compared. As noted above, disproportionality refers to one population being out of proportion with respect to the general population, demonstrated in this debate by showing that the percentage of children of one racial group in the foster care system is greater (or less) than the percentage of children 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. If the population of non-African American children in that 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 one. In this example, African American children are three times more likely to be in out-of-home care than non-African American children are.

This method is effective for revealing disproportionality and is the basis for the racial equity scorecard developed by the Casey-CSSP alliance for Racial Equity in Child Welfare.[18] 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 only the incidence of 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, "[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."[19] The point is a good one: the need to create a measure of disproportionality that can travel as a concept from one jurisdiction to another, from one time to another, or from one decision process to another. A measure that describes how disproportionate a system is in aggregate, not only how much one group is represented compared to one other group, would allow for a more direct investigation of systemic causes of disproportionality, variations in disproportionality over time, and differences in disproportionality experiences across jurisdictions.

Individual jurisdictions may have a different set of racial groups that are represented at different rates in the local foster care system and in local dependency courts. If the concept of disproportionality is limited to only one racial group or to only one group at a time, a broader, more variable, trend might be missed. 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. A consideration of only the rates per 1000 would not allow an effective comparison of disproportionality across these two counties.

The common "coefficient of variation" measure,[20] which is a useful

---

16. *Cf.* Ronald G.Thompson & Wendy F. Auslander, *Risk Factors for Alcohol and Marijuana Use among Adolescents in Foster Care*, 32 J. of Substance Abuse Treatment, 61 (2007); Jennifer Macomber, *et al. Coming of Age: Employment Outcomes for Youth Who Age Out of foster Care through Their Middle Twenties* (2008); Cheryl Zlotnick, *What Research tells us About the Intersecting Streams of Homelessness and Foster Care*, 3 Am. J. of Orthopsychiatry 319 (2009); Irene Yen, *et al. From Homeless to Hopeless and Healthless?: The Health Impacts of Housing Challenges among Former Foster Care Youth Transitioning to Adulthood in California*. 32 Issues in Comprehensive Pediatric Nursing 77 (2009).

17. Bartholet, *supra* at 911.

18. *See* Derezotes, *et al. Evaluating Multi-Systemic Efforts to Impact Disproportionality through Key Decision Points* (2008).

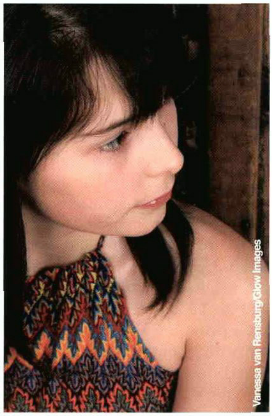
19. Shaw, *et al. Measuring Racial Disparity in Child Welfare*, 87 Child Welfare, 23, 24 (2008).

20. 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.

21. William H. Greene, *Economic Analysis* (6th ed. 2007).

form of the standard deviation,[21] may be an effective method to address the comparison. 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 is 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 is 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.

To illustrate, consider a hypothetical jurisdiction where the rate of African American children in care per 1,000 African American children is 3, the rate of White children in care is 1 per 1000, the rate of Latino children in care is 2 per 1000, the rate of Asian children is care is 0.5 per 1000, and the rate for Native Americans is 4 per 1000. In this jurisdiction, the average rate across groups is 2.1 (3 + 1 + 2 + 0.5 + 4 = 10.5 / 5 = 2.1), and the standard deviation is 1.28. Taking the two together, the coefficient of variation is 0.61 (1.28 / 2.1). Compare this to a second jurisdiction where the rate of African American children in care per 1,000 is 5, the rate of White children in care is 1 per 1000, the rate of Latino children in care is 1 per 1000, the rate of Asian children is care is 1 per 1000, and the rate for Native Americans is 5 per 1000. In this jurisdiction, the average rate is 2.6, and the standard deviation is 1.96. This produces a coefficient of variation of .75. This exhibits greater variation across rates of children in out of home care in the second jurisdiction than in the first. Knowing this difference allows for an investigation into the contributing causes of the difference.



A related interpretation issue arises from 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."[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."[23]

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 vary greatly. Some jurisdictions are able to track a wide variety of key measures while some jurisdictions must track measures by hand. To compound the difficulty, different terms and concepts are sometimes defined to mean different things in different jurisdictions. As Giovani Sartori notes, "the wider the world under investigation, the more we need conceptual tools that are able to travel."[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 need, ultimately, 'universal' categories—concepts which are applicable to any time and place,"[25] and these need to be empirically precise as to what is being measured or compared.

Probabilities

Probabilities can sometimes be misleading or suggest non-obvious conclusions. For example, if thirty percent of all children in care are African

★★★★★★★ ★★★★★★★

THE CHALLENGE WITH ESTABLISHING BROAD-BASED MEASURES OF DISPARITIES . . . IS THAT THE EXISTENCE AND QUALITY OF DATA SYSTEMS VARY GREATLY

★★★★★★★ ★★★★★★★

22. Johnson, *et al. Addressing Disproportionality and Disparity in Child Welfare: Evaluation of an Anti-Racism Training for Community Service Providers*, 31 CHILDREN AND YOUTH SERVICES REVIEW, 1 (2009); Hill, *supra*.

23. Wulczyn & Lery, *supra* at 5.

24. Giovani Sartori, *Concept Misinformation in Comparative Politics*, 64 The American political Science Review, 1033, 1034 (1970).

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.

American and seventy percent are White, does this indicate that African American children are more or less likely to be placed in foster care? The likelihood of being placed in care is conditioned by the likelihood of having a case brought before a juvenile dependency court.[26] How do these probabilities relate? What can be said 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.[27] 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.

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[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 whether the probabilities for White children and for African American children 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, the probability that a White child will be placed in out-of-home care if the case is substantiated can be calculated to be 0.61.[29] The probability that an African American child will be placed in out-of-home care if the case is substantiated can be similarly calculated to be .84. This indicates 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. Put differently, 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. In contrast, for every 10 African American children who are determined to be victims of child abuse or neglect, eight will be removed from their homes and placed in care.

The ability 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—is helpful in moving the discussion forward.

Multicollinearity

Multicollinearity describes those instances when two explanatory variables are closely correlated or often do not occur independently.[30] Discussion about the causes of disproportionality and disparities often relies upon the statistical tool of regression,[31] and one rule of regression is that there is no multicollinearity among the explanatory variables. The statistical consequence of multicollinearity is that the variances of the regression coefficients (the parameter estimates) for those explanatory variables are quite large.[32] Put another way, regression outputs are much less precise in the presence of multicollinearity, and statistical significant is less likely to be achieved.

"Whenever two supposedly independent variables are highly correlated, it will be difficult to assess their relative importance in determining some dependent variable."[33] In the words of another classic work, "Multicollinearity constitutes a threat—and often a very serious threat—both to proper specification and effective estimation" of regression models.[34]

The reason for this threat is that the regression procedure does not have enough independent variation in each of the explanatory variables to calculate their effects on the dependent variable with much confidence. Variation from one explanatory variable is used to estimate the effect of that variable only, not of any other variable. In the words of another statistician, this becomes a problem "because there would be no way of knowing whether the dependent variable variation was due to variation in the first or second variable... the common variation is ignored."[35]

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).

28. National Survey of Child and Adolescent Well-Being Research Group, *National Survey of Child and Adolescent Well-being (NSCAW); CPS Sample Component Wave 1 Data Analysis Report* (2005).

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.

30. Greene, *supra*.

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-multicollinearity, as well as normally distributed error terms, linearity of the relationships and constant variances).

32. Peter Kennedy, A Guide to Econometrics 206 (5th ed. 2003).

33. H. M. Blalock, Jr., *Correlated Independent Variables: The Problem of Multicollinearity*, 233 Social Forces 42 (1963).

34. Donald Farrar and Robert Glaube, *Multicollinearity in Regression Analysis: The Problem Revisited*, 93 The Review of Economics and Statistics 49 (1967).


Mikhail Kokhanchikov/Glow Images

A simple example might be that height and weight of shipping boxes tend to correlate—on average, taller shipping boxes tend to be heavier than shorter shipping boxes. If height and weight were both included in a regression model as explanatory, the results of the model would suffer from multicollinearity. In this case, a solution might be to remove both height and weight from the model, and use surface area (calculated from height and weight) instead.

In a similar sense, race and poverty are too closely related for both to be included as measures in a regression model. The historical context and the institutional and social legacies of race result in measures of race and poverty being correlated. A substantial body of research has demonstrated that people of color experience low wages and unemployment at disproportionate rates.[36] The problem of multicollinearity arises from this empirical correlation between race and the likelihood of experiencing poverty.

The problem of multicollinearity in this case is primarily one of interpretation. One could formulate a regression model that included race as an explanatory variable and find that race is a statistically significant predictor of the outcome variable. If the same regression model were modified to include poverty, and if the results then suggest that race is no longer a statistically significant predictor of the outcome variable, what is the appropriate conclusion?

Because race and poverty are collinear, the regression model is not properly specified resulting in very imprecise findings. The regression findings in this care suggest that both race and poverty relate to the outcome, and it is difficult to assess their relative importance. In this case, the statistical model demonstrates that it is a false dichotomy to posit that the outcome occurs because of poverty *or* because of race. The empirical analysis shows that *both* poverty and race, along with a host of other factors, are statistically related to the outcome.

The insights of another statistician are appropriate to this discussion. Ultimately, he suggests, "multicollinearity is likely to prevent the data from speaking loudly on some issues, even when all the resources of ... theory have been exhausted."[37] This is particularly apt for discussion around race and poverty. Data on race and poverty do not clearly suggest that either race or poverty is more important than the other in the analysis of disparities and disproportionality.

Conclusions

The data on race, foster care, disproportionality, and disparities indicate that something is out of balance in the juvenile dependency system, but it is the interpretation of these data that can suggest how to understand the imbalance, how to address the imbalance, and ultimately how to solve the imbalance. The juvenile dependency field would benefit from a deeper understanding of the data and tools needed to make intelligently informed decisions about disproportionality and disparities. This paper offers a critique of some uses of data and statistical tools in the debate about racial disproportionality and disparities in the child welfare system, and acknowledges that these might be but a few of many. The goal of this line of analysis is not merely to critique, but to promote a clearer understanding of the potential reliance on false logic and dangerous misunderstandings—a step toward formulating policies, practices, and tools for reducing disproportionality and disparities in a meaningful way. ★

35. Kennedy, *supra* at 207.

36. See, *inter alia*, Applied Research Center and the Center for the Study of Social Policy, *Check the Color Line, 2009 Income Report* (2009); Arloc Sherman, *Income Inequality Hits Record Levels, New CBO Data Show* (2007); Deborah Reed & Jennifer Cheng, *Racial and Ethnic Wage Gaps in the California Labor Market* (2003); and Arthur F. Jones Jr. & Daniel H. Weinberg, *The Changing Shape of the Nation's Income Distribution 1947–1998. U.S. Census Bureau* (2000).

37. Blanchard *supra* at 449.

JESSE RUSSELL
is a senior research associate at the National Council of Juvenile and Family Court Judges in Reno. (russelje@shu.edu)