

Module 3 Video 5: Ethics and Equity in Algorithms

[00:00:00] Hi, welcome back. In this video, we are going to talk about the implications for ethics and equity in your data journalism story of Algorithmic Accountability and how we talk and think about the word "fair."

[00:00:17] So what is Algorithmic Accountability and what is fair? Let's say that we're working with a school and we want to understand whether the rates that young people in that school are being disciplined are fair. Let's say that there are three different types of students in this school. There are orange, yellow and blue students at this school and a certain percentage of each of those groups of students is getting disciplined. And we want to know, are they getting disciplined in a fair way or are some groups of students getting disciplined at a greater or lesser rate that makes it unfair?

[00:01:06] There are a lot of different ways to think about this, to use data to answer that question and to tell stories about the question of is something fair? So here we see a chart that shows us the number of students in each group of students, student group one, student group two, and student group three. And the next line of our chart shows the count of the students that are disciplined. So in the student group one, there's seven out of the ten have been disciplined. In student group two five out of the 20 have been disciplined and in student group three, fifteen out of 30 have been disciplined. So the rest of the chart shows a whole number of different ways to use math or data to calculate different measures of whether or not this is fair. So there is the rate, which is simply making it a proportion out of 100. So for student group one, seventy five out of 100 students would be disciplined. In group two, twenty five out of 100 would be disciplined. And in student group 3, 50 out of 100 would be disciplined. So basically, you're taking the number of students and you're standardizing it. So what if our school had a hundred of each of these students rather than 10, 20 and 30 of different of these different types of students? What would it look like then?

[00:02:45] The second is a rate that is relative to a specific student group. So you pick a student group. In this case, the authors picked group number two and they used the basic rate, the line above, to calculate a relative rate. So relative rate is compared to a student group that we've chosen. At what proportion are these student groups being disciplined? So this is this is not a raw rate, like seventy five out of hundred, but it's a rate that compares the number of students in this group that are being disciplined to a standardized number that you choose. And the fact that you get to choose the kind of what's what's called baseline is the fastest way to embed a world view. What is normal? What do you want to standardize by. The composition index is the, another way to do this. And that is showing it's sort of the way that if people are talking about discipline, that like twenty five percent of student group A got disciplined. But they're only 10 percent of the entire school enrollment. So that's a composition index. So relative to their prevalence in the school itself, what is their prevalence within the group of students that were disciplined?

[00:04:23] And then the two final ways are the differences in composition, which are percentage points, and these are positive and negative numbers. So in other words, how much over or underrepresented is each student group in the group of students who were disciplined relative to how represented they are in the student body. So in other words, if student group one is 10 percent of the entire student body, but they are 20 percent of the entire group of people who have been disciplined. They are overrepresented by maybe ten point six percentage points. In student group two it's underrepresented by fifteen point two percentage points.

[00:05:18] And then the final possibility. I mean, these aren't. This is not an exhaustive list. There are other ways. But the final possibility that we're showing you here is the relative difference in composition of student disciplines and enrollment, which is where you combine the rates and the composition index.

[00:05:34] So all of these are different when we're talking about algorithmic accountability. These are all different algorithms or methods of calculating. Is student discipline in this school fair? So you don't actually need to know the math of all of these different ways. What you need to know is that there are lots of different ways to calculate something as simple as is the number of students in student group one that are getting disciplined, fair compared to the number of students in

student group two that there are very many different ways to answer this question, using the data. And how you choose which one of those algorithms you choose is going to highlight different experiences and going to conceal other experiences. So it's possible to include in your data story just the simple rate. This is real data from an anonymous school district. But you can see if we just showed the rate of students who experienced one suspension or more by race or ethnic group. We can see that black children in this school district have a much higher rate of suspension. What we can't tell is what any of the denominators are. So how how big is this actual problem? We can't really. We can get a sense that it's a problem, but we can't really tell how big or how fair the problem is by only looking at the rate.

[00:07:11] Relative rates, of course, compare rates of students in one or more racial ethnic groups with the rates among a selected group. Usually it's in this type of analysis, it is the group of students that you expect are being treated the best. So in this case, we are. I didn't make this, but the people that made us chose white students from an ethics and equity point of view, that is extremely problematic because that immediately sends the message that whatever is happening with white students is normal, is the baseline, and we're comparing all their students to that. And there's no way to avoid that really in a relative rate comparison. Even if you change white for one of the other groups of students, it's extremely problematic from an equity point of view. So we try and find other ways to convey more nuanced information than a straight rate without simply making it compared to X type of kids.

[00:08:18] We can compare two composition's the proportion of students who are suspended to the proportion of the students in the school, or we can try and do something like the relative difference. One of the issues here, though, of course, is that the more nuanced the data analysis gets, the harder it is to understand and the more it might be used to tell a story. That's not quite true. People who don't quite understand the data. So it's a tradeoff in data journalism whether you're going to use the most robust, nuanced kind of data analysis available to you, which will not tell as dramatic of a story sometimes and might be harder to understand. Or are you going to tell the simplest story which might actually conceal certain lived experiences? So that's an ethical and equity question that you yourself have to make in each of your data stories and how you think about fair. Read the Nick's piece on algorithmic accountability in the assigned readings and then understand that for algorithms to measure fair they have to be choosing something, like they're either optimizing the equal false negative rates, the equal false positive rates, the equal positive predictive values or statistical parity. These are four different ways among many about how to measure how fair an outcome is, almost exactly like what we've just been talking about in terms of student discipline. And the thing about an algorithm is that it can not optimize all of these things at the same time. It's not mathematically possible. So in a data analysis, in an algorithm, somebody somewhere along the line is choosing what definition of fair they want to make, the most important definition of fair. And the algorithm is based on that.