

Are Logic and Math Relevant to Social Debates?

Michael A. Lewis

The Silberman School of Social Work at Hunter College

Follow this and additional works at: <https://scholarship.claremont.edu/jhm>



Part of the [Gender and Sexuality Commons](#), [Logic and Foundations of Mathematics Commons](#), [Philosophy of Language Commons](#), and the [Race and Ethnicity Commons](#)

Recommended Citation

Michael A. Lewis, "Are Logic and Math Relevant to Social Debates?," *Journal of Humanistic Mathematics*, Volume 10 Issue 1 (January 2020), pages 383-401. DOI: 10.5642/jhummath.202001.21. Available at: <https://scholarship.claremont.edu/jhm/vol10/iss1/21>

©2020 by the authors. This work is licensed under a Creative Commons License.

JHM is an open access bi-annual journal sponsored by the Claremont Center for the Mathematical Sciences and published by the Claremont Colleges Library | ISSN 2159-8118 | <http://scholarship.claremont.edu/jhm/>

The editorial staff of JHM works hard to make sure the scholarship disseminated in JHM is accurate and upholds professional ethical guidelines. However the views and opinions expressed in each published manuscript belong exclusively to the individual contributor(s). The publisher and the editors do not endorse or accept responsibility for them. See <https://scholarship.claremont.edu/jhm/policies.html> for more information.

Are Logic and Math Relevant to Social Debates?

Cover Page Footnote

I'd like to thank Samuel Aymer, Alexis Kuerbis, Rufina Lee, Willie Tolliver, Eri Noguchi, Anna Kramarsky, Jessica Bovo, Marc Intrater, Roy Wasserman, Laurie Salitan, and Charles Reis for reviewing all or part of this article. I take full responsibility for any errors that remain.

Are Logic and Math Relevant to Social Debates?

Michael A. Lewis

The Silberman School of Social Work at Hunter College, New York City, USA
michael.a.lewis@hunter.cuny.edu

Synopsis

Social debates, as well as discussions about certain highly charged issues, such as racism, gender identity, and sexuality, usually turn on the uses or mentions of key words. That is, the conclusions we can draw from such discussions depend on how certain terms are used or mentioned in them. Yet participants in social debates may often fail to precisely define their terms or fail to make important distinctions in terms uttered by others. Both logic and mathematics pay attention to the importance of precise definitions when it comes to engaging in discussions, arguments, or proofs. Logic also makes an important distinction between the use of words and the mention of them. This article reviews ideas from mathematics and logic about definitions, use, and mention of words by applying them to central concepts which have arisen in various social debates.

Keywords: race, racism, gender identity, sexuality, intersectionality, math, logic

Race, *racism*, and *intersectionality* are terms frequently heard in public discourse. For example, members of the group Black Lives Matter are concerned that police killings of unarmed blacks are indicative of racism on the part of police forces across the country. Some whites who counter that “all lives matter” accuse members of Black Lives Matter of being racist. There are many who support affirmative action as a way of righting historical and current wrongs against “people of color” or of promoting diversity in a range of institutions. Others contend that affirmative action is “reverse racism”. A number of activists frequently use the term *intersectionality* to highlight that some people are oppressed based on multiple identities (such, as race, gender, etc.).

As I've listened to these debates, I've been struck by how logic and mathematics are relevant to them, although they rarely come up, at least not explicitly so. Many of the statements and terms I've heard in these discussions have been quite vague. I found myself wondering: what do they mean by *race*, what do they mean by *gender*, what do they mean by *intersectionality*, etc.? And, I thought, this is where logic and mathematics may be relevant. Both of these disciplines have many virtues, but an especially salient one is how they can make ideas more precise. Another is how they can assist us in drawing out the implications of our assumptions. The rest of this paper will illustrate these two virtues.

1. The Importance of Definitions

One of the things mathematicians and logicians do well is carefully define their terms, as well as draw out the implications of their definitions. Having studied these two fields to some extent, this is a practice I've also tried to adopt.

I teach at a school of social work in New York City, typically courses in statistics, public policy, and economics. These are courses, especially public policy and economics, where debates/discussions about race, racism, etc. come to the fore. Like many professions, social work has an organization that occasionally makes statements about key social issues. That organization is called The National Association of Social Workers (NASW), and in 2007 it authored a document called *Institutional Racism & The Social Work Profession: A Call to Action* [7]. On page 6 of that document, *racism* is defined as, "the practice of discrimination and prejudice based on racial classification supported by the power to enforce that prejudice." On that same page, "prejudice" is defined as, "the negative (or positive/idealized) attitudes, thoughts, and beliefs, about an entire category of people formed without full examination of the facts," and "discrimination" is defined as, "acting on the basis of prejudice."

Many of my students, as well as many activists I've heard, define *racism* similarly to the way the NASW document does. *Racism* is also frequently defined as the belief that someone is inferior or superior simply because of their race. These are two very different definitions of *racism*, and one of the easiest ways to see this is to draw out their logical implications. This can be done with the help of the following thought experiment.

Imagine we have two people, Mary and Sue. Mary defines racism as the NASW document does, while Sue uses a definition like the one in the previous paragraph. They get into a discussion about whether Tim is a racist. Tim happens to be black, and both Mary and Sue have heard him claim that blacks are superior to whites. Mary, who is also black, claims that Tim cannot be a racist because racism requires discrimination and prejudice, as well as the power to enforce such prejudice. She states that being black and, therefore, a member of an oppressed group, Tim does not have such power.

Sue objects to Mary's argument. She argues that since Tim has openly stated he believes blacks are superior to whites, and since racism is nothing but the belief that one racial group is superior to another one, Tim is clearly a racist.

As I said earlier, both mathematicians and logicians pay careful attention to definitions [12, 2]. Logicians typically distinguish among the following four types: *lexical*, *stipulative*, *rhetorical*, and *precising*.

A *lexical* definition is one where a concept is defined in terms of how it's used among the members of some community of language users. A *stipulative* definition is one where the agent defining the term states the meaning they have in mind, whether or not their definition is consistent with one used by some linguistic community. A *rhetorical* definition is one where the agent defining the term does so with the intention of getting the audience to view some aspect of the world as the definer does. A *precising* definition is one which attempts to define a term such that it becomes less ambiguous or vague. An *ambiguous* term is one with at least two meanings, and a *vague* one is a term whose meaning is unclear.

These different types of definitions are not mutually exclusive. Here is a question: what type(s) of definition is the earlier definitions of *racism* meant to be? The answer to this question isn't clear, and that may be at the heart of disputes about racism. The debate above appears to be about which definition of *racism* is the correct one. But the notion of a "correct" definition seems to apply only to lexical ones.

To see what I mean, consider Mary and Sue again. Suppose each of them presents their definition of *racism* as a lexical one. This immediately raises the question of which community of language users each has in mind. Suppose the community Mary has in mind is different from the one Sue has in mind.

It is possible for two different communities to differ about their understanding of what racism is. Suppose also that Mary's definition accurately captures the sentiment of the community she has in mind, while Sue's captures the sentiment of the one she has in mind. Then it would seem that Mary's and Sue's definitions are both correct, even though they differ from one another. That is, a correct lexical definition of a term is one which accurately captures how the relevant linguistic community defines it.

The notion of a *lexical definition* raises the question: Can there be correct or incorrect definitions in an absolute sense? That is, can there be correct or incorrect definitions which don't depend on what members of given communities believe? It isn't obvious there can be. Consider the term *racism* again. What absolute standard would we compare a given definition of *racism* to in order to determine whether it's the right one? If one discussant in a debate defines *racism* as NASW does, while another employs the alternative definition, it's hard to see how one of them could be wrong in an absolute sense.

If I'm right about this, it may make sense to simply acknowledge that different linguistic communities define terms like *racism* differently. Then whether a term would apply in any given case would depend on which community's conception we have in mind. Approaching things in this way might allow Mary and Sue to see that their disagreement about Tim is semantic, that it's difficult to know if such disagreements can be settled on the basis of some absolute standard, and that, therefore, it might be prudent not to be too quick to render a judgement about whose definition is correct.

Earlier I said a rhetorical definition is one where the agent defining the term does so with the intention of getting the audience to view some part of the world as they do. Let us relate this to the concept of *racism*. Take a look again at NASW's definition:

the practice of discrimination and prejudice based on racial classification supported by the power to enforce that prejudice

And here is the alternative one:

the belief that someone is inferior or superior simply because of their race

These may not be simply lexical definitions but rhetorical ones as well. The definition NASW uses seems to be trying to get the audience to conclude that racism is not just about belief but also about social power. That is, just *believing something* is not enough to make someone a racist. The believer must also be in a social position to enforce that belief. That is, they must be in positions to hire and fire people, make and remake the country's laws, be owners/controllers of the means of production, etc. This is why proponents of defining *racism* as NASW does often contend that blacks cannot be racists. Blacks simply do not have the requisite social power. Apparently, proponents of this definition want others to draw this conclusion as well. See, for instance, [11].

Proponents of the other definition of racism, arguably want to divorce the concept of *racism* from social power. That is, they want us to conclude that racism is distinct from who in our society does or doesn't have power. Therefore, blacks, even if we assume that they lack power in our society, can still be racists.

What we have here is an apparent debate about how to define a term being connected to the question of which racial groups have social power in our society. This is the type of thing which happens with rhetorical definitions. And, again, if it is unclear how any definition can be correct or incorrect in an absolute sense, it's also unclear how a rhetorical one can be.

Even though there may not be absolutely correct or incorrect definitions, we might want to see how a definition "works out" by applying it to particular cases in order to see if it coheres with other beliefs we hold. If application of it results in our having to renounce certain beliefs that we take to be reasonable, this *might be* a reason to abandon or refine the definition in question.

Recall that in my initial example, Mary and Sue were discussing whether Tim, a black man who believes blacks are superior to whites, is a racist. According to NASW's definition, he is not. According to the alternative one he is. Now consider the case of Steve.

Steve is a poor white man, member of the Ku Klux Klan (KKK), and believes all blacks are morally, culturally, and intellectually inferior to whites. He is also unemployed, addicted to a prescribed opiate, and spends his days and evenings frequently high, while watching reruns of *Roseanne*. Steve does not hire or fire anyone. He does not vote on any legislation.

He is not a police officer authorized by the state to use deadly force whenever he believes such force is warranted. Let us stipulate that Steve has little or no power in our society. Now here is the question: is Steve a racist?

Based on the alternative definition, he is. Based on NASW's definition he is not. This is because to be a racist, according to that definition, it's *necessary* for that person to hold certain beliefs and to possess social power. Also, according to that definition, these two conditions are *jointly sufficient* for someone to be a racist. By lacking social power, Steve can't be a racist, just as Tim can't be. But then we're left to conclude that a white male member of the KKK who believes in the inherent inferiority of black people is not a racist. I suspect most people would find this conclusion odd, to say the least. If so, this might be a reason to abandon NASW's definition of racism or refine it.

There is a possible response to what I have just said. Someone might argue that as a white man, Steve benefits from white privilege and that, therefore, he can be a racist. But neither definition says anything about white privilege. So, this response is not relevant to the discussion. Anyone who thinks benefitting from white privilege is part of what it means to be a racist would have to include this condition in their definition. Neither of the two definitions I have considered does. Again, only the second one leads us to conclude that Steve is a racist. But then we would have to conclude that Tim is too.

Even though there may be no way of determining if a definition is correct or incorrect in an absolute sense, that does not mean that "all definitions are created equal." The social positions occupied by some in the U.S. render their definitions more authoritative than others. To see what I mean, consider the United States Supreme Court.

The United States Constitution, as well as pieces of legislation, contain a lot of words and technical terms, many of them ambiguous. Yet, the nation's system of law doesn't state that anyone's definition of such terms is as good as anyone else's. Instead, that system allows the definitions of nine judges with lifetime tenure to rule the day. In 2018, in a widely watched case, the United States Supreme Court upheld President Trump's right to restrict people from certain countries, mainly majority Muslim ones, from entering the United States, a policy sometimes called the *travel ban*.

That decision is based on the court's interpretation, in a 5-4 ruling, of a law called the *Immigration and Naturalization Act*. I might define key terms of that law differently from the five justices who upheld the President's ban. Their definition, however, is decisive, not mine.

Here is another example of the crucial role played by definitions in debates about race and racism. New York City has a few specialized public high schools, which students are admitted into on the basis of their standardized test scores. For the most recent year I could find data for, more than half the students enrolled in these schools were Asians, while only about 10% were black or Latinx [9].

New York City's Mayor, Bill De Blasio, has proposed a change in the admissions process for these schools, which has turned out to be quite controversial. I do not need to go into the details of his proposal, or the controversy surrounding it, to make the point I want to make. Suffice it to say that the mayor wants to deemphasize the role played by test scores in admission into these schools. Instead, admission decisions would be determined by students' ranks in their middle schools. That is, those students ranked at the top of their middle schools would be the ones offered spots in the city's specialized high schools. De Blasio believes this change would increase access for Black and Latinx students. But his proposal has received a significant amount of criticism from some members of New York's Asian communities [13].

Those who support the Mayor's proposal believe it would be a fairer way to allocate spots in these schools. De Blasio, along with his supporters, argues that performance on the current test is affected by test preparation courses (test prep), and many low-income Black and Latinx students come from families who can't afford such preparation. Opponents argue that many Asian parents are low-income immigrants, yet they make sacrifices to send their kids to test prep, sacrifices Black and Latinx families could make too if they wanted to. De Blasio's opponents also argue that performance on an objective test is a much fairer way to make admissions decisions than relying on the more subjective criteria middle schools use to rank their students.

As I have read about this debate, I have been drawn to one question: what does it mean for an allocation process, in this context, to be *fair*? That is, what are the necessary and sufficient conditions, or, less stringently, the criteria such an allocation process must meet in order for us to say it's fair? To see how complicated this can get, consider the following.

Some have claimed, as stated above, that many low-income Asian parents make great sacrifices to send their kids to test prep, while comparable Black and Latinx families do not. Some have also argued that this is a result of cultural differences between Asians, on the one hand, and Blacks/Latinxs, on the other. The argument is that many Asian immigrants come from cultures where high-stakes testing is a far larger determinant of children's life chances and that, therefore, they are more culturally prepared for this New York City version than many Black and Latinx families are.

I do not know if this claim is true. I am no expert on Asian, or any other, immigrant cultures and how they compare to Black and Latinx ones. But suppose, for the sake of discussion, this is true. This would mean that kids raised in Asian families enjoy an advantage that those in Black and Latinx families do not, namely the *social capital* of having parents more culturally primed to make specific sacrifices on their behalf in the context of educational test prep. And this social capital advantage would be resulting in a higher average performance on the specialized admissions test and, therefore, a higher rate of attendance at these highly coveted schools. Yet kids don't choose the families/cultures they're born into. Is allocating certain benefits to *kids* on the basis of cultural differences between their *parents* fair? That depends on how we define *fair*.

My guess is that proponents and opponents of de Blasio's proposal differ in their ideas about what *fair* means in this context. It may be good to get those ideas out into the open. For one thing, if these differing conceptions were made explicit, it might turn out that there is enough agreement to work out some type of compromise. But even if that were not the case, there might be another benefit: people would have a clearer sense of what the debate is about and, therefore, be better informed if/when they chose to take sides.

2. Set Theory and Intersectionality

Another term which comes up in a number of social debates is *intersectionality* [1]. One Google definition for this term is

the interconnected nature of social categorizations such as race, class, and gender as they apply to a given individual or group, regarded as creating overlapping and interdependent systems of discrimination or disadvantage.

I think something like the above definition is what people have in mind when they use this term. Yet this definition is vague.

To see what I mean, consider the case of Enrico. Enrico is a black man who is also wheelchair-bound. He applies for a job as an information technologist but is not hired. Assume that is all we know about Enrico's case. Here is a question: was Enrico not hired because he faced intersectionality-based discrimination or disadvantage? To determine the answer, we need to know if notions like "interconnected nature of social categorizations" and "overlapping and interdependent systems of discrimination or disadvantage" apply. But in order to determine if these apply, we need to be clear about what they mean. Are we? What does it mean for social categorizations to be "interconnected in nature"? What does it mean for disadvantage or discrimination to be "overlapping and interdependent"?

In the previous section, I defined a precisifying definition as one which attempts to define a concept in order to make it less ambiguous or vague. Let me try to construct a precisifying definition of *intersectionality*. I will not pretend that my definition is the "be all and end all" of this concept. That is, it may fail to capture every nuance of what people mean when they use the term *intersectionality*. But even if my attempt to precisely define the concept is not "the last word," my purpose will have been served. That is because all I'm trying to do is illustrate the role more precise definitions can play in clearing up vagueness in the terms we use to discuss social issues. I think a more precise definition of *intersectionality* can be constructed using *set theory*.

Consider the following sets: the set of people discriminated against or facing disadvantage because of their race, the set of those discriminated against or facing disadvantage because of their gender, the set of those discriminated against or facing disadvantage because of their sexuality, and the set of those discriminated against or facing disadvantage because of a physical disability. These, of course, do not exhaust the possible sets of people who could be discriminated against or face disadvantage, since discrimination and disadvantage can occur for reasons other than race, gender, sexuality, and physical disability. Let us call any set of people discriminated against or facing disadvantage, on any basis whatsoever, a D-set ("D" for *discriminated* against or facing *disadvantage*).

Now I can propose a more precise definition of *intersectionality*:

an individual is discriminated against on an intersectional basis if and only if they are a member of at least two D-sets.

So, in order for the concept of *intersectionality* to apply to the type of discrimination or disadvantage faced by someone, they *must* be a member of at least two D-sets (necessity) and this is *all* that's required (sufficiency). So, going back to Enrico's case and if we think of sets, in this context, as groups of people, "all" we have to do to see if he faces intersectional discrimination or disadvantage is determine if he's a member of at least two D-groups. "All" is in quotes because I do not want to suggest that it is easy to determine if someone faces discrimination or disadvantage, especially since in many cases discrimination or deliberately disadvantaging someone is illegal and, therefore, those who engage in such behavior typically try to hide it. Also, there is a certain degree of uncertainty when it comes to the nature of the discrimination or disadvantage faced by members of more than one D-set.

To see what I mean, consider Mary, Joe, and Tim. Suppose Mary is a middle-income straight black woman, Joe is a middle-income straight white man, and Tim is a middle-income straight black man. Arguably, all else equal, Mary would be the one among these three who would face intersectional discrimination or disadvantage. That's because she's a member of two D-sets: women and blacks. So, in some sense, Mary would face *more* discrimination or disadvantage than would Joe, or Tim. But *more* here could mean one of two things: it could mean *additively more* or *interactively* (as in *interaction effect*, a term from statistics/quantitative social science) *more*.

To see the distinction I am making, suppose there are ratio scales for measuring the degrees of discrimination or disadvantage someone faces because of their race, gender identity, sexuality, social class, etc. Suppose further that these scales can be combined to obtain an overall degree of discrimination or disadvantage faced by someone. Let R be the measure of discrimination and disadvantage based on race and G the measure based on gender identity. The higher R or G is, the more discrimination or disadvantage one faces. So, Tim, who's black, would presumably have a higher R value than Joe would. Let us assume Mary and Tim have the same R value but that her G value is higher than Joe's and Tim's, both of which are 0. To get at what folks may have in mind when they use the term *intersectionality*, we need to think about how G and R might enter a person's overall discrimination or disadvantage equation (let us call this their D-equation).

Does a person's equation take a form something like this:

$$D = G + R?$$

Or something like this:

$$D = G + R + GR?$$

The first equation above is additive while the second is interactive. That is, in the first equation, the effect of G on a person's level of discrimination or disadvantage is constant (and equal to 1), and the same is true of R 's effect. In the second equation, the effect of G on Person's level of discrimination or disadvantage depends on the level of R , and R 's effect depends on the level of G . This can be seen by calculating the first partial derivatives of the person's D -equation with respect to G and R , respectively:

$$\frac{\partial D}{\partial G} = 1 + R, \quad \text{and} \quad \frac{\partial D}{\partial R} = 1 + G.$$

To see how all this plays out, consider the following numerical examples.

Suppose the additive versions of Joe's D -equation is

$$D_{Joe} = G + R = 0 + 0 = 0,$$

Tim's D -equation is

$$D_{Tim} = G + R = 0 + 5 = 5,$$

and Mary's D -equation is

$$D_{Mary} = G + R = 5 + 5 = 10.$$

And assume the interactive versions of Joe's D -equation is

$$D_{Joe} = G + R + GR = 0 + 0 + 0 \times 0 = 0,$$

Tim's D -equation is

$$D_{Tim} = G + R + GR = 0 + 5 + 0 \times 5 = 5,$$

and Mary's D-equation is

$$D_{Mary} = G + R + GR = 5 + 5 + 5 \times 5 = 10 + 25 = 35.$$

It's clear that Mary, due to intersectionality, suffers more from discrimination and disadvantage than do Joe or Tim, regardless of whether we use the additive or interactive model. But whether intersectionality operates additively or interactively makes a pretty big difference, since 35 is more than 3 times as large as 10.

Obtaining data on ratio level measures of various types of discrimination or disadvantage is a challenge, to put it mildly. This renders mathematics a limited tool when it comes to defining the concept of *intersectionality*. But that does not mean our exercise has been completely useless. If nothing else, it's allowed us to think more precisely about how intersectionality might operate to compound the discrimination or disadvantage faced by members of some groups in society.

Mathematics, specifically set theory, can also be applied to understand a notion that frequently came up after Barack Obama was elected president: the idea of a post-racial America. This is the idea that the United States is no longer a place where people face discrimination on the basis of race [14]. According to this perspective, there may still be other types of discrimination, but racial discrimination isn't one of them. From a mathematical point of view, this amounts to the claim, with U.S. residents being the universal set, that the set of people discriminated against because of their race is the empty set. That is, if the U.S. really is a post-racial nation, the D-set made up of people discriminated against on the basis of race has no members whatsoever. I will leave it to readers to decide if they think this particular D-set is really empty.

3. Gender Identity and Fuzzy Set Theory

The website for the National Center for Transgender Equality [4] states that, "most people—including transgender people—are either male or female. But some don't neatly fit into the categories of "man" or "woman" or "male" or "female." The site goes on to tell us that some people identify as a gender which blends male and female, some identify with a gender that is neither male nor female, and some don't identify with any gender at all.

As I've read about discussions and debates about transgender and non-binary persons, I've almost immediately started thinking about set theory, as well as the limitations of mathematics as a tool for modeling complicated realities.

When I discussed intersectionality in the previous section, I had in mind standard set theory. Standard set theory is also called *crisp set theory* [10]. Crisp set theory has one distinguishing trait: an entity in a particular universe of discourse is either in a given set or not. For example, consider the set of positive integers $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \dots\}$ where the universe of discourse is the set of all numbers. Either a number is an element of the set of positive integers or it is not. There are no “in-betweens.” The number 25 is an element of this set, while the number -25 is not. An objection many transgender and non-binary persons have raised might be stated as the problem of others trying to impose the tools of crisp set theory on them.

In terms of crisp set theory, with human beings the universe of discourse, there are only two genders: male and female. That is, we have two sets: males = {all human beings who're male} and females = {all human beings who're female}. According to the traditional view, a person's gender places them in one *and only one* of these two sets. And which of the two a person's gender places them in depends on that person's anatomy. That is, which of the two sets one's gender places them in isn't something they choose—it is chosen for them.

What the National Center for Transgender Equality is telling us is that this way of seeing things simply doesn't reflect the reality of how people actually understand their genders. In a paper appearing in this Journal, Andres Sanchez [8] applies crisp set theory to both gender identity and sexuality. If, like Sanchez, we want to remain within the confines of crisp set theory but want to use it to more realistically model how people identify themselves, we need more crisp sets; here are four possibilities:

- 1) males = {all human beings who *identify* as male}
- 2) females = {all human beings who *identify* as female}
- 3) neither gender = {all human beings who *identify* with a gender which is neither male nor female} and
- 4) {all human beings who do not *identify* with any gender at all}.

Notice the emphasized terms in the sets above. These are intended to highlight the challenge that transgender and non-binary persons are posing to the more traditional way of understanding gender. Gender isn't something to be imposed from "outside" but instead depends on how people identify themselves.

The National Center for Transgender Equality website also hints at another challenge to the traditional way of understanding gender. Notice that it also mentions that some people identify as both male and female. That is, some identify as a blend of these two genders. If we want to represent such blending with the tools of set theory, we must leave the confines of crisp set theory. A version of set theory which may be useful here is *fuzzy set theory*.

Fuzzy set theory is a tool which was developed to model *vagueness*, the lack of sharp boundaries between one thing and another. Vagueness is ubiquitous in the real world. Is an adult who is 5'10" tall or short? That depends on where we draw the line between tall and short. Perhaps someone 3'0" is clearly short, while someone 7'0" is clearly tall. But what about that 5'10" adult? Fuzzy set theory was developed to address just these types of situations.

The key idea from fuzzy set theory is the notion of a *membership function*. A membership function is what allows one to specify the degree to which an element of the universe of discourse falls within some set of interest. The domain of such a function is the universe of discourse relevant to the set we're trying to construct. Going back to gender identity, the domain would be the set of all human beings. The range of a membership function is the closed interval $[0, 1]$. Regarding gender identity, the range of the membership function would amount to stating that a person's gender identity can take any real number value between 0 and 1 inclusive. We could assign a 0 to those who identify as male and a 1 to those who identify as female. Those who identify as a blend of male and female, or both male and female, could pick or be assigned some number other than 0 or 1.

Fuzzy set theory may do a better job than crisp set theory at representing how people in the real world identify their genders. But it has a key limitation. For those who identify as blends of male and female, we do not know the membership function for assigning degrees of maleness or femaleness. That is, we do not know how to assign real numbers, between 0 and 1 inclusive, reflecting the degree to which people identify as male or female.

Even with this limitation, though, fuzzy set theory may capture “the spirit” of how some people identify better than crisp set theory does.

4. The Use-Mention Distinction

In November of 2018 after a screening of his new movie *The Green Book*, actor Viggo Mortensen waded into a bitter controversy by saying “*N-word*.” What he said was, “for instance, no one says *N-word* anymore.” Those who attended the event, including Mortensen himself, have stated that the context for his comment was his claim that even though many whites now refrain from using that word, that does not mean they do not still hold racist views. So, Mortensen didn’t appear to be calling anyone a “*N-word*.” But that didn’t stop others, including his co-star Mahershala Ali, from criticizing him for saying it, although Ali did accept Mortensen’s apology for doing so. Many blacks debate whether they should say the word. But I think it is fair to say that almost all blacks contend that whites, no matter how well-meaning, should never say it [3]. This is dangerous territory to move into, but even here I think logic may help clarify what’s at issue.

One of the key distinctions made in logic is the *use-mention* distinction [2]. To see the difference between use and mention, consider the following two sentences:

Sentence 1: Michael Jackson was a member of The Jackson Five.

Sentence 2: “Michael Jackson” was a member of The Jackson Five.

The first sentence is true, while the second one has never been true. In the first sentence, the word “Michael Jackson” refers to a person who died on June 25, 2009. In the second one, “Michael Jackson” refers to the name of the person who died on June 25, 2009. The *person* Michael Jackson was, during an early period in his life, a member of the Jackson Five. The *name* “Michael Jackson” never was because names cannot be members of singing groups.

Logicians would say that the word “Michael Jackson” is *used* in the first sentence but only *mentioned* in the second one. Words are used when they are meant to refer to things in the world or, in the case of Michael Jackson since he is now dead, things that were in the world. Words are mentioned

when they're meant to refer only to themselves. When words are mentioned in sentences, as opposed to used in them, logicians typically place those words between quotation marks, just as I did above. Now consider the following two sentences:

Sentence 1: Michael Jackson is a *N-word*.

Sentence 2: Michael Jackson was probably at some point in his life called a "*N-word*."

In the first sentence, "*N-word*" is used to refer to Michael Jackson. In the second one, the word is being mentioned as a term that others may have used to refer to Michael Jackson at some point in his life. Notice that it is being used not to refer to Michael Jackson but to refer to itself as a word that others may have used, at some point, to refer to Michael Jackson.

To complicate things a bit more, "*N-word*" can be used in different ways. It can be used in a derogatory way, say, by a Neo-Nazi or Klansman. Or it can be used, some would say "reclaimed," by black people to refer to members of that community not in a derogatory way but as a kind of term of endearment. This is a use often associated with hip-hop culture.

As the debate about "N-word" has unfolded, I haven't seen much explicit attention paid to the use-mention distinction. But that distinction is relevant. Here's a quote from Mortensen indicating how this distinction often goes unrecognized:

"In making the point that many people casually used the 'N' word at the time in which the movie's story takes place, in 1962, I used the full word. . . ." [3]

But arguably Mortensen did not use the "N-word"—he only mentioned it. Even though the use-mention distinction often goes unrecognized in this context, it is one that I am sure people could appreciate if it were brought to their attention. It is easy to imagine how that might happen.

Suppose Mortensen had said, "my co-star Mahershala Ali is a "N-word"." What he did say, which clearly was not this, was controversial enough. Had he used "N-word" this way to refer to his co-star, I suspect the reaction and

criticism would have been much harsher. It is likely that Ali would not have been so quick to accept his apology, and Mortensen's career would've been ruined beyond repair.

What I have said about "N-word" arguably applies to other words which have come up in similar debates. Examples are "C-word" and "Q-word".¹ Some contend that "C-word" should never be used by anyone, while others claim that it is okay for women to use it [6]. There are those who state that "F-word" or "D-word" shouldn't be used by those who are not gay, although some who identify as gay contend that it's acceptable for members of that community to use them [5]. The use-mention distinction is relevant here too. There is a difference between 2a and 2b, as well as 3a and 3b:

Sentence 2a. Anya is a "C-word".

Sentence 2b. Anya has probably been called a "C-word" at some point in her life.

Sentence 3a. Klaus is a "F-word".

Sentence 3b. Klaus has probably been called a "F-word" at some point in his life.

Also, some contend that these words can be used in derogatory or reclaiming ways. Recognizing the distinction between using and mentioning these terms, as well as the different types of uses of them, we can see how complex these debates really are. Hopefully, we will get to a society where everyone can agree that none of these terms should ever be used in derogatory ways. But perhaps reasonable people can disagree about whether they should ever be reclaimed or merely mentioned.

To anticipate a possible objection, I am not suggesting that use or mention of all the above terms is the same in every respect. Perhaps differences in racial, gender, and sexuality-based oppression are such that whether anyone

¹Just as many find saying the actual word, instead of "N-word", offensive some find saying certain words used to refer to women or gay persons offensive. That's why above I said "C-word", for a derogatory term sometimes used to refer to women, and "F-word", "D-word", and "Q-word" for terms that are sometimes used in a derogatory way to refer to, respectively, gay men, gay women, and gay persons in general.

should use or mention any one of them varies depending on the term we have in mind. For example, perhaps the country's history of racial oppression is such that no one should ever use or mention "N-word," but under some circumstances it may be okay to use, in a reclaiming way, or mention the others. My goal here is not to take a position on any of these issues—it is to show that perhaps logic's use-mention distinction may help clarify them.

Acknowledgements: I'd like to thank Samuel Aymer, Alexis Kuerbis, Rufina Lee, Willie Tolliver, Eri Noguchi, Anna Kramarsky, Jessica Bovo, Marc Intrater, Roy Wasserman, Laurie Salitan, and Charles Reis for reviewing all or part of this article. I take full responsibility for any errors that remain.

References

- [1] Kimberlé Williams Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics", *University of Chicago Legal Forum*, University of Chicago Law School (1989), pages 139–168. Available at <https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8/>, last accessed on January 27, 2020.
- [2] Harrie de Swart, *Philosophical and Mathematical Logic*, Springer, 2018.
- [3] Scott Feinberg and Gregg Kilday, "Viggo Mortensen apologizes for using n-word during 'green book' panel", *The Hollywood Reporter*, November 8, 2018. Available at <https://www.hollywoodreporter.com/race/viggo-mortensen-says-sorry-using-n-word-q-a-1159813>, last accessed on January 27, 2020.
- [4] National Center for Transgender Equality, *Understanding non-binary people: How to be respectful and supportive*, Technical report, 2018. Available at <https://transequality.org/issues/resources/understanding-non-binary-people-how-to-be-respectful-and-supportive>, last accessed on January 27, 2020.
- [5] Randall Jensen, "Should the LGBT community use words like 'faggot' or 'tranny'? Yes, call me a faggot," *Huffpost*, March 16, 2012; updated December 6, 2017. Available at https://www.huffpost.com/entry/call-me-a-faggot_b_1296886, last accessed on January 27, 2020.

- [6] Kayleigh Llewellyn, “Reclaiming the c-word”, *Pypo*. Available at <https://www.pypo.com/reclaiming-the-c-word/>, last accessed on January 27, 2020.
- [7] National Association of Social Workers, *Institutional racism and the social work profession: A call to action*, Technical report, National Association of Social Workers, 2007. Available at <https://ncwwi.org/index.php/component/k2/item/1259-institutional-racism-and-the-social-work-profession-a-call-to-action>, last accessed on January 27, 2020.
- [8] Andres Sanchez, “My sets and sexuality”, *Journal of Humanistic Mathematics*, Volume 8 Issue 1 (January 2018), pages 359–370. Available at <https://scholarship.claremont.edu/jhm/vol8/iss1/18/>, last accessed on January 27, 2020.
- [9] Grace Segers, “A guide to the controversy around NYC’s specialized high schools”, *City & State New York*, June 7, 2018. Available at <https://www.cityandstateny.com/articles/policy/education/new-york-city-specialized-high-schools-admissions-guide.html>, last accessed on January 27, 2020.
- [10] Michael Smithson and Jay Verkuilen, *Fuzzy Set Theory: Applications in the Social Sciences*, Sage Publishing, 2006.
- [11] Mzwakali Sobantu, “Black people can’t be racist”, *Pambazuka News: Voices for Freedom and Justice*, October 13, 2015. Available at <https://www.pambazuka.org/governance/black-people-can-t-be-racist>, last accessed on January 27, 2020.
- [12] Daniel Velleman, *How to Prove it: a Structured Approach*, Cambridge University Press, 1994.
- [13] Hyo-Jung Yoon, “Asian community pushes back against elite schools plan”, *Kings County Politics*, June 5, 2018. Available at <https://www.kingscountypolitics.com/asian-community-pushes-back-against-elite-schools-admission-plan/>, last accessed on January 27, 2020.
- [14] Joy Zarembka, “Still Waiting for a ‘Post-Racial’ America”, *Institute for Policy Studies*, December 4, 2008. Available at https://ips-dc.org/still_waiting_for_a_post-racial_america/, last accessed on January 27, 2020.