The Fundamental Attribution Error, Microaggressions, Harassment, and Gaslighting of Transgender Athletes: A Bi-Directional Analysis of the Role of Stereotypes in Transmisogynist Harassment in Single-Gender Sports

Christina Friedlaender
Rachel McKinnon

For inclusion in Microaggressions and Philosophy (Routledge), Lauren Freeman and Jeanine Weekes Schroer, eds.. Please cite published article, when possible.

[Abstract]

I. Introduction

This paper is concerned with forms of microaggressive harassment targeting trans women athletes. In particular, we offer a bi-directional analysis on the role stereotypes play in a perpetrator’s unintentional use of the fundamental attribution error (FAE) by examining this error as a precursor to transmisogynist microaggressive harassment in the context of single-gender sports.¹ In this case, the role of fundamental attribution errors and misperception provides the perpetrator with faulty moral justification for engaging in microaggressive harassment. By examining microaggressive harassment through the use of FAE, we show why epistemic gaslighting is central concern for understanding the harms microaggressions cause.

Much of the contemporary research and discussion about microaggressions predominately frames agents as unintentionally and unconsciously committing these harmful acts despite consciously holding so-called progressive and egalitarian values. Since implicit bias is taken as the primary precursor to microaggressions, descriptive and normative analyses focus on the absence of intent to harm. This agent-failure model centers a definition of microaggressions wherein an agent’s unintentional failure to avoid causing harm happens despite the agent’s conscious desires or value commitments to avoid doing so. Unlike intentional and conscious macroaggressions, this framing treats unintentional and unconscious microaggressions as necessarily attributionally ambiguous, making them difficult identify for perpetrators, bystanders, and targets.

By contrast, intentional microaggressions are largely undertheorized. The predominate model presents intentional microaggressions as a deviation insofar as they do not necessarily share the precursor of implicit bias or the feature of attributional ambiguity. However, by disconnecting intentional microaggressions from attributional ambiguity, the standard framing cannot adequately distinguish between intentional microaggressions and macroaggressions. On this model, unambiguous actions, such as the intentional display of a hate symbol or use of a slur, qualify as intentional microaggressions, even though such actions would be better

¹ While we focus on single-gender sports, we are confident that these same phenomena exist in multi-gender sports.
categorized as form of macroaggressive harassment. The relevant different here is covert versus overt behavior, as Bonilla Silva argues in the case of racism (2014). It would be odd and mistaken to classify overt behavior as a microaggression.

This paper challenges the standard framing for both intentional and unintentional microaggressions, focusing on their connections to harassment, stereotypes, and epistemic gaslighting. Regardless of type, attributional ambiguity is the core feature of microaggressions. By connecting microaggressions to harassment, we challenge the belief that harassment is largely macroaggressive. Since harassment is necessarily intentional, microaggressive harassment exploits this feature of attributional ambiguity. Like its unintentional counterpart, microaggressive harassment can result from unconscious cognitive or perceptual failures in the perpetrator. However, the agent-failure model cannot capture microaggressive harassment, or intentional microaggressions more broadly.

We argue for two additional models – desire-to-harm and failure-to-disavow-harm – that more adequately capture the relationship between intention and action when taken in conjunction with the agent-failure model. For these two models, the agent has not merely failed to avoid causing harm despite consciously holding value commitments that regard such harm as morally wrong and/or consciously desiring to avoid causing such harms. Rather, the agent intends to harm or approves of their unintentional harm ex post facto and consciously regards the harm as justified. Any account of microaggressions that prioritizes or solely relies on the agent-failure model cannot capture intentional microaggressions and mistakenly treats them as a deviation. This is broadly concerning for descriptive and normative research that takes this standard framing as its foundation.

II. Fundamental Attribution Error and Microaggressions: A Case Study in Elite Cycling
Elite bicycle racing isn’t for the faint-hearted. There are many kinds of racing: individual time trials, team time trials, long road races, criteriums, track cycling on a velodrome, inner city ‘alley cat’ races, and more. We’re going to focus, for our purposes, on criterium racing. These are typically timed events: a race may be planned for 60 minutes, and after the first few laps, the organizers project how many laps the race will take, and then start counting down laps. By switching to counting down laps, races are typically a little shorter or longer than the planned time (in this case 60min). But we’re only talking two or three minutes in either direction.

Criteriums (or ‘crits’) are like NASCAR on bikes. They typically happen in downtown settings with courses ranging from, on average, two to three city blocks of various route shapes. They’re generally some form of ‘Tetris’ block shape. Some are a simple rectangle around two blocks; some are nearly a figure-eight but the route doesn’t cross itself; some are an ‘L’ shape; and others are more complicated shapes, particularly depending on the available road directions. Some crits have a hill, others are pancake flat.

We encourage you to find a local race and watch the pro races: you’ll typically see racers going around 90 degree corners, four riders wide, at 30mph or faster. The biggest races are held just after sunset with a party atmosphere for spectators. And there are crashes. There’s a saying in bike racing: it’s not if you’ll crash, it’s when. Every pro racer has crashed at least once, usually
not their own fault. Rachel was in one race, for example, where she was rounding a corner in a large pack only to find a huge pile of riders already on the ground in front of her and nowhere for her to avoid it: you just have to hit your brakes as hard as you can and get ready to fall.

Highly skilled riders are both less likely to cause a crash themselves (but it still happens in the heat of the moment) and are more skilled at avoiding crashes as they happen or are about to happen. In pro racing, the pack is racing very tightly together: there are lots of elbows exchanged, slight leaning on other riders to avoid falling, being pushed into other riders and having to adapt, and even the rare bodycheck. A skilled rider can take, for example, a bodycheck and stay upright (and thus avoiding a crash involving other riders as well). They can also monitor the riders around them to anticipate sketchy maneuvers, and avoid bad situations that could result in a crash. This might mean a shuffle to the side where the rider uses her hips to shift the bike to the side to avoid a hazard, which is safer and more effective than steering with her hands. However, less skilled riders often receive small amounts of contact and crash, usually blaming it on another rider rather than understanding limits in their own bikehandling skills. With that in mind, here are our two cases.

Alison is a professional racer. She races lots of ‘road’ disciplines including time trials, road races, criteriums, and stage races that combine each of these. She’s a multi-year state champion and competes at the highest domestic levels, including some international competition. She races on her own, without a team, but has many chances to win. And she’s been quite successful regionally, winning a number of high-profile professional races.

Sheila is another professional racer who is highly ranked nationally and races regularly on the professional circuit. She rides for a professional team. In her role, she takes care of the team’s sprinter, making sure she’s in a good position for the finish of the races. Sheila is a strong rider, but doesn’t have the chance to win many races. Sheila and Alison have raced with each other many times.

Both Alison and Sheila are in a prestigious professional criterium. There is significant money on the line (for bike racing, at least): $1500 for first place. There have already been two crashes in the race, which is fairly normal for a tight downtown course as they’re on. Midway through the race, Alison is following a less experienced racer, who makes an unpredictable move to the left. Because everyone is racing so close, Alison has to react or the other racer might cause Alison to crash. So Alison makes a quick shift to the left to avoid the danger. In doing so, she lightly bumps the rider to her left, but it doesn’t affect the other rider’s overall position. The rider shoots a scowl at Alison, but they keep racing. Two laps later, the same thing happens to Sheila, who also adjusts appropriately. She moves to the right, causing her to lightly bump the rider to her right, who shoots her a scowl, but they keep riding without problem.

---

2 Don’t let the ‘professional’ moniker suggest that women are often paid a salary: they’re not. There is no minimum salary for women set by the Union Cycliste Internationale (UCI), the international cycling governing body. There is a minimum for men, though. A good documentary is Half the Road by Katherine Bertine (2014).
Unfortunately, people react differently to Alison’s movement than Sheila’s. A significant number of racers perceive Alison’s evasive move as “dangerous” and that she doesn’t belong in women’s professional racing. People like Alison are highly stereotyped as inherently “aggressive,” and her movement is attributed to her “aggression” rather than her avoiding another rider’s sudden movement. However, the very same people view Sheila’s move positively: shit happens in bike races. Everyone has to make a move like that from time to time. Rubbing is racing, after all. Sometimes you bump people and it’s not your fault.

Here’s the problem: Alison is widely known to be transgender. She’s even a little famous for it, often doing interviews. And since she’s successful, her participation in the sport (while fully supported by the sport’s governing body, even at the highest levels) is controversial. Many people have, unsuccessfully, complained to have her banned from competing with women. There’s a persistent, widely held stereotype that trans women are ‘really’ men, and since men are aggressive, trans women are aggressive. Even though her maneuver is unremarkable, every time she makes it, it’s attributed to her being, and thus serves as confirmation that she’s, aggressive and dangerous.

Sheila, though, is cisgender and white. She’s a team player and reasonably well-liked. She doesn’t make waves, and there are no stereotypes about who she is. Her identical maneuver, under sufficiently identical circumstances is differentially attributed to circumstance rather than inherent personality traits. We have a classic case of the fundamental attribution error. In this case, it’s due to Alison’s gender identity (as trans), and the stereotype that trans women are ‘aggressive.’ We now have our case.

III. Fundamental Attributions Error and Microaggressions
Our fundamental concerns are how this fundamental attribution error constitutes a precursor to transmisogynist microaggressions and, in particular, the role this error plays in microaggressive harassment and epistemic gaslighting. Within the psychology literature, microaggressions are predominately defined as: “brief and commonplace daily verbal, behavioral, or environmental indignities, whether intentional or unintentional, that communicate hostile, derogatory, or negative slights and insults” towards members of oppressed groups, such as people of color and LGBTQ+ individuals (Sue et al. 2007). Both the academic and popular culture conversation about microaggressions have heavily centered uninterrupted microaggressions (UMSs).

One explanation for this is the unique moral problems arising from the connection between implicit bias as a precursor to microaggressions, given agents’ lack of intentionality and conscious awareness (Hall and Payne 2010; Zheng 2016; Friedlaender 2018). Centering implicit bias as the precursor to microaggressions raises a particular kind of moral concern for those who consciously hold so-called progressive or egalitarian values, such as supporting trans rights, because a person can harbor unconscious biases and engage in unintentional microaggressions that violate their consciously held values (Sue 2010). For instance, when such a person is made aware of their implicit biases, they often express an alienation between their unconscious

---

3 For a discussion of some stereotypes of trans women (and the attendant phenomenon of stereotype threat), see McKinnon (2014).
attitudes and their conscious beliefs by accepting they harbor these biases and by disavowing them as “not who they really are,” i.e., as unrepresentative of their consciously endorsing values and beliefs (Glasgow 2016). By centering implicit bias in UMS, research on microaggressions operates on what we call an agent-failure model: an agent commits microaggressive harms despite their conscious commitment to avoid doing so.

The fundamental attribution error, however, reveals a different precursor to microaggressions. In this case, the attribution of aggression to Alison’s gender identity as transgender is a cognitive error based off a set of sex/gender stereotypes. Crucially, those guilty of this error do not exhibit or express an alienation between unconscious behavior and conscious values or beliefs. Instead, bystanders perceive Alison and Shelia’s behaviors differently based off of consciously held beliefs about the relationship between biology, gender identity, and gendered stereotypes about character traits.

In order to commit this FAE, one must be consciously committed to a worldview in which: 1) the discrete binary of biological sex categorization, i.e., male or female, accurately reflects all humans; 2) there is a necessary connection between this binary sex categorization and one’s gender identity and/or expression, such that trans, intersex, gender nonconforming, genderqueer, and nonbinary individuals are regarded as medical anomalies to be fixed upon birth, or as psychiatric anomalies deemed mentally ill, lying, or rejecting their biological essence in a way that is “unnatural;” and finally, 3) that this necessary connection produces a set of gendered stereotypes that are regarded as generally stable and reliable, consciously or unconsciously.

Thus, when bystanders see Alison and Sheila exhibit the same behavior but attribute it to different causes, their conscious beliefs about sex and gender alter their perception of the behavior unconsciously. This is a kind of top-down cognition (Palmer 1999, 84-85; Andrews 2001). Shelia’s behavior isn’t attributed to her gender identity because she is a cis woman, who, due to her biological sex categorization, is less likely to be aggressive than those assigned to the opposing category of ‘male’ at birth. Alison’s behavior, on the other hand, is attributed to her gender identity because she is perceived as being biological male, and men are stereotyped as being aggressive. If those making this mistake were made aware of this difference in attribution, it is unlikely they would disavow the unconscious error because it does not necessarily conflict with their conscious values and beliefs.

This kind of fundamental attribution error has a bi-directional effect. Because the governing stereotype of trans women is that they’re ‘men,’ and since men are stereotyped as aggressive, people expect to see aggression from trans women athletes. This has a bottom-up effect via the fundamental attribution error: noticing a behavior from Alison that might count as aggressive, but wrongly attributing it to Alison’s identity rather than chalking it up to

---

4 This isn’t to say that attributions of her behaviors aren’t causally connected to her perceived gender identity; however, in a heavily cissexist society, people’s gender identity qua cisgender doesn’t rise to the level of the typical perceiver’s conscious awareness. Sheila is merely perceived as a woman, not a cis woman, even though she is a cis woman.
circumstances; whereas, had Sheila performed the same behavior, the observer would have been inclined to chalk it up to circumstances and not her character or fundamental nature.  

The fundamental attribution error also has a top-down effect in that Alison’s neutral behaviors will be wrongly, but sincerely, perceived as aggressive. The likelihood of this error can be exacerbated by hostile attribution bias wherein bystanders attribute hostile intent to otherwise neutral or ambiguous behavior. Moreover, the presence of confirmation bias reinforces the worldview that supports such stereotypes because ensures that an bystanders initial misperception of neutral behavior or misattribution of aggressive behavior counts as further evidence for it. Stereotypes are thus \textit{self-reinforcing}, with the FAE as one mechanism.

This bi-directional effect is a cognitive failure in two instances: wrongful attribution and misperception. Similar to the mechanisms of implicit bias, the individual is not aware or in control of making a fundamental attribution error. However, the fundamental attribution error is different than the standard model of implicit bias as the precursor to microaggressions. First, there isn’t a disconnect between the wrongful attribution or misperception and the bystander’s conscious beliefs and value judgments. In this case, the bystander is already consciously committed to stereotypes about biological sex categories and trans women. The wrongful attribution couldn’t occur but for these conscious, background commitments. The misperception that follows is pernicious much in the way that ignorance can be pernicious.

Kristie Dotson defines \textit{pernicious ignorance} as “any reliable ignorance that, in a given context, harms another person (or set of persons),” wherein reliable ignorance is “ignorance that is \textit{consistent} or follows from a predictable epistemic gap in cognitive resources.” (Dotson 2011, 238). \textit{Pernicious misperception} is thus a misperception that harms others due to predictable cognitive errors that follow from one’s consciously endorsed stereotypes. Pernicious misperception follows from pernicious ignorance, particularly ignorance regarding stereotypes.

Furthermore, because the precursor for microaggressions is generally understood to be implicit bias, the fundamental attribution error changes the role of intentionality when such errors result in a microaggression. The role of intentionality in microaggressions has generally been regarded as binary: either intentionality is present or it isn’t. Intentional microaggressions receive little treatment in the literature because they are currently defined in ways that make them indistinguishable from microaggressions.

For instance, Derald Wing Sue defines intentional microaggressions as microassaults: a sub-set of microaggressions defined by the presence of intentionality and the overtness of their meaning (Sue 2010, 28-31). His examples include: the use of racial slurs, jokes based off of sexist stereotypes, and the display of hate symbols such as the Confederate flag. Sue argues that microassaults are similar to ‘old fashioned’ forms of racism, sexism, or heterosexism insofar as they are intentional, overt, and regarded as publically condemnable behavior. Unintentional microaggressions, on the other hand, are largely defined by the presence of attributional

\footnote{While we have analytically separated top-down and bottom-up processing for explanatory process, the two do not function as wholly separate modes of processing. For further discussion, see Rauss and Pourtois (2013).}
ambiguity. Attributional ambiguity occurs when the causes or motivations of one person’s behavior towards another are unclear, such that the witnesses and the target(s) might second-guess whether the behavior was, in fact, a microaggression.

One example, from McKinnon (2014), is a trans woman being selected for additional TSA screening at an airport. Was she singled out because she was ‘clocked’ for being trans, and TSA agents tend to subject visibly trans people to invasive additional screening? Or was she truly randomly selected? It’s unclear to the trans woman, and this ambiguity in attributing the reason for being singled out will naturally cause her to suspect the more negative of the options. This tends to carry with it a spike in social anxiety, feeling less safe, and other negative outcomes.6

If a fundamental attribution error is a different precursor to microaggressive behavior, then this offers a new way of considering the role of intentionality and attributional ambiguity. The binary approach to intentionality requires that some microaggressions lack attributional ambiguity. But as we argue, macroaggressions are overt and thus lack attributional ambiguity. Without this feature, there is no way to distinguish intentional microaggressions from macroaggressions, which is a necessary step in conceptualizing harassment as microaggressive, not just macroaggressive. By taking attributional ambiguity as a necessary feature of all microaggressions, we are in a better position to capture the relationship between the fundamental attribution error, microaggressions, and microaggressive harassment.

Intentionality ought not be understood as a binary but rather as a spectrum. The fundamental attribution error doesn’t fully lack intentionality because it arises from an agent’s conscious beliefs, i.e., beliefs about the truth of transmisogynist stereotypes, even if the moment of error lacks consciousness awareness. While the presence of stereotypes might have an impact on unconscious cognitive processes, these stereotypes differ from implicit bias because they are consciously endorsed by the agent. This conscious endorsement impacts the agent’s ability to realize their error because their perception is already altered by the stereotypes. Again, the presence of confirmation bias in this case only serves to further distort the bystanders’ perception of what took place by providing them with further “evidence” to endorse the stereotype.

By starting from attributional ambiguity as a necessary condition, we challenge the emphasis on the agent-failure model within microaggression research. As it stands, the predominate moral concern is about well-intentioned agents whose implicit bias produces unconscious behaviors that violate their conscious value commitments.7 If most microaggressions were the result of agent-failure, then one might expect that pointing out the behavior and explaining why it is likely a microaggression would frequently result in an apology

__________________________

6 AA sources.
7 This might also help explain the phenomenon of ‘white fragility’ (and its intersectional siblings) discussed by DiAngelo (2011, 2018). When confronted with racist behaviors, the conversation usually quickly shifts to how the agent didn’t intend the harm they caused. The implication is that they are less morally culpable for the harm. In practice, this is usually deployed (often unintentionally) as an exonerating excuse: “Isn’t it worse if someone meant the harm?” they ask rhetorically.
and a commitment to changing the behavior from that moment on. Indeed, this is usually sufficient to remedy the harm.

However, anyone who has tried to point out microaggressive behavior knows that this rarely happens. The perpetrator often denies that they have done anything wrong and attempts to provide an alternative explanation for the behavior. This attempt at an alternative explanation is only possible because microaggressions are attributionally ambiguous. Alternative explanations are not possible if the behavior is overt, such as the use of a transmisogynist slur. The agent-failure model cannot explain why those criticizing Alison’s racing as dangerous will likely deny they have committed a fundamental attribution error and sincerely believe their alternative explanation for any behavior that results from the error. And this sincerity is important: because the FAE is a top-down cognitive process, it affects the agent at the perceptual level. However, bad actors also intentionally trade on the attributional ambiguity of microaggressive behaviors as a way to harm with impunity.

More interestingly, even when the agent recognizes they have committed a microaggression, they are less likely to disavow the behavior in that they express a sense of alienation from it. It is more likely that the agent will blame the target for being “overly-sensitive” or claim that the target is “playing the victim” (Manne 2018). In these cases, the agent might also tell the target that they ought to focus on the “real” problems of transmisogyny, i.e., macroaggressions.

This kind of behavior suggests that even if agents do not intentionally commit a microaggression, the agent-failure model is inadequate to capture what has happened. On an agent-failure model, the agent must both intend not to harm and disapprove of the harms they have committed; otherwise there is no conflict between their unconscious behavior and their conscious values. But this rarely happens. It is more likely that the person will be ambivalent about whether they actually caused the harm—rather, they’ll think that the target is overreacting or being uncharitable—failing to disapprove of their harmful action and its cause. In the case of the FAE, the agent might even approve of the harm they have caused and instead regard it as justified, not recognizing that the FAE has played a role in their behavior. For instance, if Sheila commits this fundamental attribution error during race, sees Alison’s behavior as “dangerous” to her, and engages in harmful behavior as a result, then Sheila is also likely to try to justify that harm as a form of self-defense or retribution for Alison’s “dangerous” behavior.

Regardless of whether the microaggression is acknowledged by the perpetrator, the FAE leads to a different kind of analysis of microaggressions that more realistically captures the frustrating and often harmful dynamics between agents, targets, and bystanders. People often react poorly to having their microaggressive actions pointed out to them. They often deflect the harm as being “minor” or defend themselves with an alternative explanation for their behavior, which happens to trade on the attributional ambiguity of the situation. They may do so because they already endorse the stereotypes present. The problem with stereotypes is that if you genuinely believe to be true, then perceptually relying on them, consciously or not, is not considered a harm. More concerning, the role of stereotypes in the FAE is also often used to justify the harm they have already caused. This leads to microaggressive harassment and epistemic gaslighting – both of which constitute further harms.
IV. Microaggressive Harassment
The relative lack of academic work on microaggressive harassment reflects broader limitations with the agent-failure model. If microaggressions are understood in terms of unintentional implicit bias, then microaggressions cannot be a form of harassment. Harassment requires intentionality. Understanding FAE as a different precursor to microaggressions both broadens the types of actions that constitute microaggressions and the types of actions that constitute harassment.

Harassment is largely regarded as overt, conscious, and intentional. It lacks attributional ambiguity. When harassment is macroaggressive, the harasser cannot attempt to offer an alternative explanation for their behavior and can only contest the terms of harassment when caught. For instance, if a bystander repeatedly yells transmisogynist slurs during the race, then that person might attempt to argue that the word is not actually a slur and/or that transmisogyny is not real. However, this response differs from the kind of alternative explanation found in attributionally ambiguous behavior. For starters, if the harasser tries to avoid responsibility by denying transmisogyny, the target will not second-guess whether they have actually been harassed. The action and its meaning isn’t ambiguous, even if the harasser denies the features of social reality that make this action and its meaning unambiguous.

But this framing does not capture the full range of harassment. Harassment can operate more insidiously. Microaggressive harassment occurs when the behavior is attributionally ambiguous, leaving room for targets and bystanders to second-guess what has actually occurred. This kind of harassment is harder for targets to point out and address because the harasser can, necessarily, always give an alternative explanation, one that might be seen as reasonable even to the target. In this case, the alternative explanation centers on the cause of the behavior. Causal ambiguity can even leave targets second-guessing the meaning of the action.

In this case of microaggressive harassment, this problem is two-fold because there becomes an open question about whether harassment has occurred and what kind harassment it is. It becomes easier for harassers to avoid detection and it becomes more likely that targets second-guess their own experiences. Moreover, this kind of harassment imposes a new kind of burden on the target: the person must prove that the alternative explanation is wrong, that a microaggression has, in fact, occurred, and that the harasser was acting intentionally.

This is an impossible task. When the target of microaggressive harassment calls out the behavior, they are generally expected to prove that the harasser holds the stereotypes in question, has committed an (often unconscious) cognitive error, and acted directly as the result of that error in an intentional way. But people can’t read the mental states of others. When this kind of proof is asked of the target, they are doomed to fail. This failure is then used to justify more harassment, punish the target for making “false accusations,” and further diminish the epistemic credence given toward their testimony. In the case of microaggression harassment, the alternative

---

8 For instance, under U.S. law, harassment is defined as a form of repeated or continuously unwanted contact in which the sole purpose of the behavior is to annoy, intimidate, or cause emotional distress in the target. It is taken to be an intentional form of torment.
explanation does not necessarily question the existence of transmisogyny but rather calls into doubt whether a particular action is transmisogynist. The problem with this kind of doubt is unique because the target is liable to internalize it as well.

The FAE illuminates just how insidious microaggressive harassment can be. If Shelia or others wrongly perceives Alison’s maneuver as being aggressive due to her gender identity and sees her as dangerous, Shelia will not merely fail to see her error and the problematic stereotypes it presupposes; she will feel justified in acting on it. In this case, microaggressive harassment happens when Sheila acts in ways that are harmful to Alison but sees those harms as a justified response because she genuinely believes her perception of Alison as dangerous. This sort of harassment can range from social bullying to institutional mobbing (McKinnon forthcoming a). Shelia might exhibit covert anti-social behavior towards Alison and encourage other riders to do so. This sort of “mean girls” bullying and mobbing is ambiguous because when done effectively, Alison might be left second-guessing whether this bullying is actually happening or if it just so happens that other riders are increasingly unfriendly.9 This can damage Alison’s trust broadly within the sport because she never really knows who is genuinely trustworthy or who is pretending to be trustworthy in hopes of finding way to further damage Alison’s reputation in the sport. A natural response to this is for Alison to become less trusting in others, and perhaps beginning to socially isolate herself. For example, she may warm-up alone, not socialize pre/post-race, and not introduce herself to people to make new friends.

Moreover, this kind of microaggressive harassment often leads to more overt, macroaggressive harassment. Shelia might report Alison’s behavior to the relevant institutional authorities in hopes of calling into question the legitimacy of Alison’s participation in the sport. This can further increase microaggressions directed at Alison from other players and cycling officials. People may ‘reason,’ that even if Alison can prove all of the claims wrong, well, there must be something she’s doing wrong if all these people are complaining about her: where there’s smoke, there must be fire, right? No: sometimes it’s just bullies blowing smoke.

**V. Harassment and Epistemic Gaslighting**

Whether intentional or not, the success of deflecting the blame for committing a microaggression, especially resulting from the FAE, by exploiting the attributional ambiguity of microaggressions dovetails with *epistemic gaslighting*. Briefly, gaslighting takes two forms. First, it can take a form of psychological warfare, where the agent attempts to convince the target that the target is crazy and can’t trust their own perceptions and must therefore depend more heavily on the agent. This is always intentional, even if the agent doesn’t have the concept of ‘gaslighting.’ Second, it can take a subtler epistemic form where an audience to a victim’s testimony of harm isn’t given appropriate credibility, where the audience inappropriately doubts the victim is capable of perceiving events reliably. “Directly, or indirectly, then, gaslighting involves expressing doubts that the harm or injustice that the speaker is testifying to really happened as the speaker claims” (McKinnon 2017, 168).

---

The epistemic form of gaslighting (hereafter just referred to as ‘gaslighting’ for simplicity’s sake) often takes the form of responses to victim testimony such as: “I know [abuser], they wouldn’t do something like that,” “You’re on edge, so you expect to see [abuse] everywhere,” “This isn’t that serious, you’re overreacting,” and so on. But they all trade on the victim telling someone about the harm they suffered, and the audience thinking that the victim isn’t reliably perceiving things, and so shouldn’t be believed.

Now recall typical responses to a victim claiming that they have suffered a microaggression. “I’m sure that’s not what [agent] meant, you’re just reading into things,” “You’re overreacting, it’s not that serious,” “You think everyone is out to get you,” and so on. Classic gaslighting. The audiences of the claims that the victim suffered a microaggression are effectively saying that the victim isn’t reliably perceiving things. An important implication of this is that sometimes (perhaps even often) the victims begin to doubt themselves: their self-trust erodes (Jones 1996, 2001, 2002, 2010; Zagzebski 2012; McKinnon 2017, forthcoming a).

This erosion often carries forward with the victim to other contexts. This can result in a tendency in the victim to doubt their own perceptions each new time they think that they experienced a microaggression. They will also be less likely to tell someone due to this erosion of self-trust. “Did that really happen? Maybe I am overreacting. No one else saw it.” And this happens even when a microaggression certainly did happen.

This is where the concept and practice of providing a sanity check to the victim is critically important (Sue et al., 2008; Friedlaender 2018). A sanity check is when a bystander witnesses a microaggression (or even a macroaggression) and somehow affirms to the victim that the bystander witnessed the act too. This can be verbal: “Did that really just happen?” asks the victim. “Oh yeah, totally,” replies the bystander.

Here is an example. Veronica was at a conference presentation on decision theory and on a topic that Veronica had spent considerable time researching, nearly writing a dissertation on the topic years before. She was up-to-date on the literature. The speaker gave his talk. Veronica noticed that the speaker made a crucial assumption at the beginning of the talk that his interlocutor would never grant, and that this assumption seemed to drive the conclusion more than any of his arguments. So she asked about this.

The speaker replied that he did not understand the question. So Veronica rephrased. He still didn’t understand. She tried one last time. Nothing. Not wanting to take up additional question period time, she gave up. Matthew asked the next question, wanting to directly follow-up on Veronica’s question. He phrases his question almost identically to Veronica’s phrasing (any differences were unimportant). Suddenly the speaker understood and gave an answer, although one that was unconvincing to Veronica and Matthew.

---

10 There’s a saying in marginalized communities that, “You’re not crazy if they’re actually out to get you.”
Veronica immediately supposed that she had suffered a microaggression from the speaker. He didn’t ‘understand’ the question when she, a woman, asked it, but he suddenly understood it when Matthew, a man, asked the very same question. This is related to what Rebecca Solnit calls ‘mansplaining’ (2008, 2014). It’s a form of epistemic injustice where we don’t give a speaker appropriate credibility; in this case, the speaker didn’t appropriately think that Veronica was asking a sensible question. But she wasn’t sure: there was some characteristic self-doubt. “Maybe I wasn’t as clear as I thought. I thought Matthew asked exactly the same question as me. Maybe I’m overreacting?”

After the session, Veronica approached Matthew and introduced herself. She asked whether what she thought happened—she suffered a gendered microaggression—and Matthew immediately responded, “Oh yeah, that definitely just happened. I’ve never seen something so stark.” Veronica and Matthew became fast friends, bonding over this experience. But Matthew just provided Veronica with a sanity check: no, she wasn’t imagining it. Yes it really happened just as she thought. Now, rather than stewing in self-doubt about whether she actually suffered the microaggression, she could have a happy cathartic conversation with a new friend. A net positive outcome!

Sanity checks don’t need to be conversational. While they’re powerful after the fact, they’re more powerful when they happen immediately in the moment of the microaggression. While speaking up may not be safe or possible due to social norms, something as simple as purposeful eye contact and a subtle knowing nod, or mouthing ‘What the fuck?!’ goes a very long way. Sanity checks can be silent, subversive. Providing quick sanity checks for victims of microaggressions is one of the more powerful things a witness to the behavior can do.

Unfortunately, things often don’t go so well in response to a microaggression. When a victim is gaslit by someone they perceive as an ‘ally’ or supporter, trust in that person—and others more generally—also erodes. If the person we trust most to believe us in a time of need doesn’t, then that is experienced as a betrayal. And if we’re betrayed by someone we trust, then we’re less likely to trust others. We keep our suffering to ourselves. Rachel argues in a forthcoming paper that habitual gaslighting can cause Complex Post-Traumatic Stress Disorder (C-PTSD) (forthcoming, a). Since a typical response to claims of microaggressions seems to be gaslighting, then it’s a small step to suggest that enough microaggressions could cause C-PTSD as well. While microaggressions are typically conceived of as ‘papercuts,’ enough papercuts will bleed a person dry.

Gaslighting thus compounds the harms of microaggressions. While the microaggression may not be particularly harmful in isolation, when the victim tries to speak about their harm, being gaslit can often be worse than the initial harm from the microaggression. This can also trigger what Rachel has called the epistemic injustice circle of hell (McKinnon 2017). Being gaslit hurts, and can cause the victim to be more emotional than they were before. So we can have a situation where a speaker is gaslit, often by invoking ‘emotion’ as the reason not to believe them, which makes them more emotional, which is taken as further evidence not to believe them, which makes them more emotional, … and so on in a vicious circle.
The problem of gaslighting and microaggressions is further compounded when we consider Lauren Freeman and Heather Stewart’s newly proposed taxonomy on microaggressions (2018). This taxonomy begins from centering the target, not the perpetrator, in accounting for types of microaggressions within clinical medicine and their harms. In this case, microaggressions can be emotional, epistemic, and existential (i.e., related to one’s identity). Outside of a medical context, emotional and epistemic microaggressions further compound the impact of epistemic gaslighting in our case. Epistemic microaggressions occur when the target’s claims about their experience are ignored, dismissed, or given less than the appropriate amount of credulity. In contrast with bystanders who provide sanity checks, epistemic microaggressions further compound the process of gaslighting. Victims who speak out about microaggressions are likely to encounter further epistemic microaggressions. In a similar vein, emotional microaggressions happen when onlookers regard the victim’s emotional responses as inappropriate, in which stereotypes concerning race and gender can further determine the “appropriateness” of a particular person’s emotional response. The target’s emotions are not taken as a sign of true distress or harm but instead actively dismissed by bystanders. Thus, when the victim tries to speak out about harm, they will likely encounter further epistemic microaggressions and emotional microaggressions as part of being gaslight by bystanders.

There is one final concern about gaslighting and microaggressions with the FAE. Frequently, perpetrators will wrongfully co-opt the language of victimhood and gaslighting when the target or a bystander calls out their behavior. That is, there is an attempt to re-position the aggressor as the victim and the victim as the aggressor. In Down Girl, Kate Manne argues that the position of being the victim is transformed when the victim depicts themselves as such (2018, chap. 7). Rather than being cast as a passive victim, the active victim claims their victimhood, or places themselves “at the center of the story” (Manne 2018, 225). An active victim centers their agency rather than the agency of their victimizer.

But this can change bystanders’ perception of the victim. If a victim can self-advocate, then how much of a victim are they really? Perhaps they are exaggerating the harm done to them or trying to become the center of attention. Perhaps they are simply playing the victim. On Manne’s account, those who are member of historically or currently oppressed groups are much more likely to be seen as playing rather than being a victim. This is because those in the oppressed position are regarded as givers not takers. The victim is demanding something from those regard the victim as a subordinate, as one who gives not takes. The active victim is thus “stepping out of line” or violating the terms of their subordinate role.

When someone is seen as playing the victim and accusing someone with more power than them within an oppressive hierarchy, that person is more likely to encounter: gaslighting, victim-blaming, a careless investigation of the harm and/or have evidence destroyed, accusations of selfishness and aggression, harassment and belittement, and a minimization or dismissal of any charges they bring against their victimizer (Manne 2018, 237–8). In our case, when Alison calls out the microaggressions of cis-women, she is likely to be regarded as playing the victim. If Alison is seen as playing the victim, then those she calls out are more likely to be regarded as actual victims, even if they take an active role in their supposed “victimization.” Alison has, effectively, “stepped out of line.”
Herein lies the concern: in order to call out the microaggression, Alison must tell the perpetrator(s) and bystanders that their perception is wrong. After all, it is their cognitive error that resulted in pernicious misperception and, in turn, microaggressive behavior. But when people like Shelia and bystanders who support her are told their perception is wrong and their behavior is harmful, they are likely to respond by co-opting the language of gaslighting. From their perspective, they are incorrectly being told that their experience is false, which, in turn, probably increases their sense that they must “protect” themselves or other people. They engage in further micro- and macro-aggressive harassment because they regarded it as morally justified self-defense. Bystanders will likely side with the harassers because they have managed to invert the victim/victimizer status. Thus, Alison will encounter gaslighting behavior that treats her as the “real” victimizer.

VI. Challenging the Conceptual Framework

Our project thus far has been to explain the particular mechanisms of transmisogynist microaggressive harassment without the area of single-gender sports by challenging the prominent model for microaggressions. The mechanism of harassments so described are not unique to single-gender sports but can be seen throughout any space which demands and polices the discrete binary of biological sex categorization. We’d like to end with a few prescriptions for future research on the definition and taxonomy of microaggressions as well as solutions to the problem of microaggressions.

First, the standard descriptive and evaluative treatment of unintentional microaggressions is nearly indistinguishable from macroaggressions. It is not surprising that researchers focus so heavily on unintentional cases because the presence of intentionally marks intentional microaggressions as deviations – deviations that are not fraught with the problem of attributional ambiguity and are not of moral concern for the agent-failure model. Because attributional ambiguity has been a sufficient but not necessary condition in defining microaggressions, the scope of the term becomes uncomfortably wide, including behaviors that more intuitively look like macroaggressions. Thus, we ought to take attributional ambiguity as a necessary feature of microaggressions. This feature provides a clear way by which we can delineate between the small slights, intentional or not, that become increasingly harmful in their accumulation but are individually hard to pin down and the larger, more overt behaviors that potentially cause more harm in a single instance and are also often illegal. Microaggressions are defined by their attributional ambiguity. Macroaggressions are not. Taking attributional ambiguity as a necessary feature also places more focus on the way microaggressions can cause the target to second-guess its cause, which raises the possibility gaslighting and makes it uniquely harmful in this case. Thus, attributional ambiguity captures a distinct feature of microaggressions, namely, that the behavior in question can be explained away in terms of a non-biased cause, which further compounds the experience of indeterminacy.

Second, starting from attributional ambiguity as a necessary condition thus also challenges the emphasis on the agent-failure model within microaggression research. On an agent-failure model, the agent must both intend to not harm and disapprove of harms they have committed; otherwise there is no conflict between their unconscious behavior and their conscious values. However, in certain cases, agents will not intend to cause harm but fail to disapprove of
the harm once they recognize their behavioral as harmful. Such agents might now approve of the harm after the fact or fail to express disapproval by trying to downplay the harm itself, i.e., by turning against the target of the microaggression. Let’s call this the failure-to-disavow-harm model. Moreover, there are cases in agents exploit attributional ambiguity of microaggressions in order to intentionally harm the target. This is best captured on a desire-to-harm model. By exploiting the feature of ambiguity, agents are able to intentionally harm another more covertly and thereby are likelier to avoid moral reproach from bystanders. If the target points out the harm, the agents can easily throw the burden of proof on the target, requiring that they prove that the behavior was microaggressive and that the particular microaggression in question was, in fact, harmed. These three models provide a more adequate scope to categorize the relationship between intentionality and microaggressions. De-centralizing the agent-failure model shifts the how we address prevention and moral evaluation of microaggressions. Since the predominate moral narrative is one in which the agent is alienated from their behavior, microaggressions research emphasizes a picture of well-intentioned agents who have accidently committed a harm. Prevention is thus a concern for how people match their unconscious bad behavior to their unconscious good values. But this the wrong picture to emphasize. The predominate problem with microaggressions is seems to be those that fall under the failure-to-disavow-harm and desire-to-harm models.

Finally, by rethinking the role of intentionally, we have provided a different precursor for microaggressions. While the emphasis on implicit bias is not surprising given the unique moral concerns that follow, it is a mistake to regard implicit bias as the only or predominate precursor. By focusing on the case of the fundamental attribution error, we hope that future research will not only include this precursor in its models but also aim to discover other precursors. Research on microaggressions is still in its infancy and broadening the potential ways microaggressions can arise and function is beneficial, especially if our ultimate goal is to prevent microaggression and their harm.

Works Cited


Freeman, Lauren, and Heather Stewart. (Forthcoming) “Microaggressions in Clinical Medicine.” The Kennedy Institute of Ethics Journal.


