

Unforgivable

Suppose there's a debate about the character of a public figure. Supporters will usually marshal a long list of positives. But detractors are more likely to present *one horrifying fact*. A fact horrifying enough to get onlookers to shake their heads and say, "Unforgivable." If this rhetorical tactic works, the detractors instantly win the debate. If you've done one unforgivable thing, you're a villain – no matter how else you spent your life.

In the Eighties, the top unforgivable offense was a Nazi past. I mean that literally. If someone could prove that you ever belonged to the Nazi Party, your name was forever mud. When the Cold War was still ongoing, the U.S. put Austrian President Kurt Waldheim on a war criminal watchlist – and almost every country in the world declared him "persona non grata." The only acceptable defense was (and remains?) age; former Hitler Youth could still rise to high places – even the papacy!

Other unforgivable offenses during this era included explicit anti-black racism, severe child abuse, and selling secrets to the Soviets. And perhaps Satanism, too. When people wanted to discredit Iron Maiden or Ozzy Osbourne, they'd accuse them of devil-worship. Unforgivable to many parents, though that rarely deterred their kids!

Nowadays, as you may have noticed, the list of unforgivable offenses has grown much longer. See J.K. Rowling.

What's the best way to understand this "Unforgivable Heuristic"? Economic theorists will likely gravitate to a signaling model. If a single damning fact conclusively confirms your "bad type," then it makes sense to diligently hunt for such facts – and publicize them when you find them. This isn't always crazy. If you discover that a teacher murdered one child, he shouldn't work with children – even if his performance was otherwise exemplary. The same goes for the Auschwitz commandant – if you ran a death camp, it makes little difference what else you did with your life.

The main problem with the signaling analysis, however, is that a large share of "unforgivable" offenses are trivial. People get ostracized for "unforgivable" tweets every day, and most of those tweets plainly do far less harm than, say, punching a stranger in a bar brawl, or cheating on your wife. The sheer randomness is also striking. When the media says, "Famous person X said Y!" they almost never bother to ask, "How many other people also said Y today?" If saying Y actually revealed definitive information, you would try to find all the Y-sayers and ostracize the whole lot of them, not join the dog pile of the day.

What's the alternative to the signaling story? Hysteria and herding. Most allegedly "unforgivable" offenses are basically uninformative. But virtually all allegedly "unforgivable" offenses make at least a few people temporarily but intensely angry.

Usually the anger just fizzles out, but if the few temporarily-but-intensely angry people are well-connected – or lucky – other people join the herd until it hits critical mass and explodes. Most herd members probably barely care about the original offense, but when a cruelty party starts, they rush to join the festivities.

The lesson: If someone appeals to unforgivability to instantly win a debate, they're not necessarily wrong. Once in a long while, a single offense really is ultra-informative. Most of the time, however, the Unforgivable Heuristic is way off-base. The single damning thing says a lot about the pettiness and conformity of the accusing, but next to nothing about the character of the accused.