

Philosophy 311: Knowledge and Justification

Fall 2013, Northern Illinois University

Geoff Pynn

Handout 9: Gendler on Stereotype Threat

The last six weeks have concerned *individual* epistemology. For the next few weeks we turn to *social* epistemology. This week and next: what is the epistemic significance of stereotypes? Third week: what is the epistemic significance of disagreement?

Gendler's paper concerns the epistemic effects of living in a society with significant racial distinctions. The question we'll discuss today is:

Do racial stereotypes affect our knowledge, even if they are disavowed?

Three preliminary distinctions:

1. *Stereotypes*. Stereotypes are generalizations about members of social categories. The stereotypes we'll be interested in are *value-laden*; i.e., the generalizations they involve value judgments. E.g.: "college professors tend to be over-educated, nerdy, liberal, unkempt, socially awkward." Value-laden stereotypes can be *negative* or *positive*.
2. *Dominant vs. aversive racism, explicit vs. implicit bias*. Racism is most straightforwardly thought of as a set of beliefs: a racist believes that certain people are superior or inferior because of their race. Racist beliefs (purport to) justify biased treatment of members of different racial groups. But in recent decades people have started to talk about a different sort of racism. *Dominant* racism involves holding racist beliefs, and bias based on racist beliefs we'll call *explicit* bias. Someone who disavows racist beliefs can still behave in ways that express bias towards members of certain racial groups. This is called *aversive racism*, and the bias expressed by aversive racism is called *implicit bias*. Some examples of aversive racism / implicit bias (both mentioned on p. 44 of Gendler's paper):
 - a. Resume effect
 - b. Ambiguous object associations
3. *Belief vs. alief*. Gendler coined the term "alief" to describe a cognitive state that is functionally similar to belief, but need not be consciously accessible or reflectively endorsed. Official definition: an innate or habitual propensity to respond to a particular stimulus in a particular way. Some examples where my beliefs can clash with my aliefs:
 - a. Glass walkway projecting over the Grand Canyon. I *believe* that the walkway will hold me, but I *alieve* that I am in grave danger.
 - b. Horror movies. I *believe* that there is no such person as Jason, but I *alieve* that Jason is about to brutally murder that innocent victim.
 - c. Poop-shaped cookies. I *believe* that the cookie is made of chocolate, but I *alieve* that it's made of feces.

Gendler's overall project: use the concept of alief to illuminate the phenomena of aversive racism and implicit bias. The proposal in section 4: racial stereotypes, even when actively disbelieved, are still powerful at the level of alief, and have significant epistemic effects.

Stereotype threat. Hypothesis: by activating your awareness of membership in a group associated with a stereotype before asking you to perform some task, you make it more likely that your performance will conform to that stereotype.

This hypothesis has been confirmed in many studies over the last twenty years.

(For a very readable overview of the idea by the person who first had it, see Claude Steele's book *Whistling Vivaldi*.)

Steele and Aronson 1995: black and white students performed comparably on the verbal GRE when told that the test was not intended to evaluate verbal ability, but black students performed significantly worse than white students when they were told that it was a test of verbal ability (see p. 799 for precise wording).

Aronson et al. 1999: white male students performed worse on math GRE when told that Asian students tend to outperform white students on math tests than they did when they were not given this information (see p. 37 for complete text).

Subsequent work shows that negative stereotypes needn't be *believed* by subjects in order for the effect to occur. They just need to be *salient*.

Mechanisms that explain stereotype threat are not well understood. Three independent factors seem to be involved:

- a. Increased anxiety ("Oh no I'm going to mess up! I'm going to mess up!")
- b. Increased self-monitoring ("Am I messing up? Am I messing up?")
- c. Increased self-regulation ("Don't think you're going to mess up!")

Each of these factors is independently known to decrease performance

What does this have to do with epistemology?

1. The mere existence of negative stereotypes about a social category affects our knowledge in two ways: it can make it more difficult to access, and it can undermine knowledge by undermining confidence.

2. Stereotype threat provides a concrete example of how the distinction between reasons for belief and causes of belief breaks down. Consider Norman the Clairvoyant: he's (supposedly) not justified in his belief because he doesn't have a *good reason* for it, even though it was *caused* in him by a reliable process. Or take someone who fails to know something because of some physical impairment, as opposed to someone who fails to know something because she is believing for a bad reason. Stereotype threat cases suggest that knowledge and justified belief involve factors that are more than merely causal but are less than explicitly rational --- what Gendler calls *easons*.