

Agustin Esteva

Professor Juan Qian

SOSC 13110 Social Science Inquiry: Formal Theory

December 12 2024

A Summary of Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A., 2016. “Stereotypes.” *The Quarterly Journal of Economics* 131(4)

“Stereotypes” (2016)[1] attempts to model stereotypes in a decision maker. Bordalo, Coffman, Gennaioli, and Shleifer present two credible results on the effects of stereotypes in a decision maker, including the idea that stereotypes are *context dependent*—it matters which groups you are comparing when forming stereotypes. Their second conclusion is that stereotypical decisions include a “*kernel of truth*.” That is, stereotypes are grounded in true differences between groups, but these difference are usually hyperbolized. To start, the paper provides an overview on the subject by introducing us to various approaches to modeling stereotypes with a few examples, as well as introduces notation for this model. Afterwards, they introduce their model by giving a mathematical description of different kinds of stereotypes for a group and then go on to discuss their model’s limitations. The third section talks about properties of their models by providing mathematics propositions about it and then comparing it to actual statistics. They then apply this model to two common stereotypes, one about women in mathematics and the second about disease testing. The paper finished by providing lab evidence and by gathering empirical evidence on political stereotypes, both which support their model.

Three main assumptions help build the model:

1. Stereotypes are context dependent. In order to make a valuable model, it must somehow include a comparison of a trait (crime, height, redhair) between groups (Black and White population, Dutch vs rest of the world, Irish vs rest of the world). “Stereotypes” takes a frequentist approach to their model, comparing groups by measuring the probability of a trait in some population against the probability of the same trait in another population.

2. Stereotypes contain a “kernel of truth.” Decision makers don’t come up with stereotypes out of thin air, they have some basis of empirical facts which they can then stretch out into a stereotype. This is the reason the model takes a frequentist approach to stereotypes— the model takes in actual data and then hyperbolizes certain aspects of it to make the stereotype.
3. People are bad at making decisions. They only remember certain things when making decisions. Usually, they remember the most striking relative difference between groups. If someone were to ask an American what the state with the oldest population is, their mind would immediately jump to Florida. While the age distribution of Floridians is similar to that of the entire country, the relative proportion of old Floridians is higher than any other state. Thus, our mind unproportionately weights Floridian’s age when thinking of this question. To address this memory recall issue, the model more heavily weights traits with higher relative proportions in groups.

Certain assumptions are made in order to simplify the model:

1. Traits of interest are fully understood by a decision maker before making decisions. Moreover, the decision maker is informed about these trait’s probability distributions. That is, decision makers are aware of the empirical truth about any trait given any population.
2. Decision makers know which two groups to compare. In our Florida example, was it clear to compare the sunshine state to every other state, or just southern states, or just compare it to Maine? This is important, because Maine actually has an even higher median age than Florida does, and thus the decision maker might not exaggerate this trait for Floridians so much [2]. In the model, the comparison group is simplified to simply be the ‘universe’ without the actual group. In our case, the comparison group would clearly be all the states except for Florida (and thus the “universe” would be the USA).

While these are all limitations of the model, they cannot be critiques of it, since they are baked into its definition. The author’s themselves concede some shortcomings of their model. To begin,

their model's account of memory recall is limited to only remembering groups by their most striking differences, but some decision makers recall traits by their likelihood in a population instead of by their representativeness. By the first assumption above, this is a very likely scenario— if a decision maker understands the whole picture with all the probability distributions, then they are more likely to know the traits with the highest likelihood in a group, and are thus able to recall them more easily. This shortcoming leads nicely into a property of the model:

Proposition 0.1. *If the group has a similar frequency of traits as its comparison (Florida's age distribution has roughly the same distribution as the country as a whole), then the representative trait chosen by the decision maker (the stereotype) is unlikely to be found in either group.*

In our example, being an old is (relatively) unlikely for either Florida or the United States. This phenomenon generates inaccurate stereotypes, and the “kernel of truth” hypothesis is strained, but not broken, since it is factually true that Floridan's have a higher relative proportion of older people.

Proposition 0.2. *In contrast, if the group has very different frequency of traits as its comparison, then the stereotypical trait is a very common trait in the group.*

That is, thinking that Dutch are tall is not an inaccurate stereotype, since the probability distribution of a tall person from Netherlands is very different than that of (for example) the United States. Another property of the model, to do with the weighting function embedded in the definition of the stereotype is as follows:

Proposition 0.3. *Suppose we can order the traits of interest in some way. If the relative frequency of the traits within a group increase as the traits increase, then decision makers associate this group with higher traits. A similar result holds if the frequency is decreasing.*

To understand this property, the article provides the example of the stereotype that women are worse than men at math. Here, the traits of interest are the SAT's math section scores. This can obviously be ordered. Data says that men are twice as likely to have a perfect math score as

women are, and their frequency of higher scores (when compared to the whole group), increase as the scores increase. By the proposition, we see that a decision maker can be understood to have the stereotype which associates men with higher math SAT scores. However, their distributions are actually very similar, and thus by Proposition 0.1, the stereotype is inaccurate. Thus, we can make sense of how decision makers overweight specific differences within groups, even though these differences are not diagnostic of the group as a whole.

The paper claims that this model accounts for the fact that stereotypes are context dependent and that it contains some “kernel of truth.”

1. To provide evidence for the claim that stereotypes are context dependent (i.e, they depend on which two groups the decision maker is comparing), the researchers create a laboratory experiment which consists of two groups, the control and the treatment group.

- The control is shown two groups. The first group is of a cartoon picture of boys, of which half are wearing a green shirt and half are wearing a purple shirt. The second group is of cartoon picture of girls, who’s colored shirt distribution matches that of the boys exactly. Given that both colors are representative of either gender, the model predicts that the stereotype will be correct, that is, people’s expectation of what color shirt girls wear will equal the actual color of shirts that the girls are wearing (same for the boys).
- The treatment is shown two groups. The first group is the same group of boys as in the control, but green shirts have become blue shirts. The second group is the exact same group of girls as in the control, with the same shirt-color distribution as in the control. The representative shirt for girls becomes green and that of boys becomes blue, and thus the model predicts that the stereotype will become that girls wear, on average, green more frequently and boys wear, on average, blue more frequently. Of course, this is not true since the frequency of girls wearing green is still one half and the frequency of boys wearing blue is one half.

Although the results contain much variation within them, the paper states that “we interpret our results, . . . , as an important proof of concept: the presence of representative types biases ex post assessment” (Bordalo, 1781). That is to say, decision makers form stereotypes in the presence of distinctive differences within groups.

2. To provide evidence for the “kernel of truth” claim, the paper uses data sets on political preferences to see if the believed positions for both groups are systematically exaggerated. The key finding of this study is that, in line with the model, stereotypes can be predicted using the average responses. That is, true means of both groups predict believed means for each group, which evidently supports the “kernel of truth” hypothesis. Moreover, the more representative a group’s extreme types, the greater the bias in beliefs about that group.

Thus, the model is supported by the evidence. Stereotypes are found to be context dependent as well as found to contain some “kernel of truth.”

While some shortcomings of the model have already been discussed above, we can now go more in depth after seeing the properties and evidence of the model. A model is supposed to be simple but still representative of the phenomenon at hand, but we argue that his model is much too simple in the way that their assumptions about context dependence are too limited. It was touched on above that it is assumed decision makers know which two groups to compare, but nowhere in the paper is it discussed the baseline assumptions about the decision maker’s inherent stereotypes before looking at the data. That is, while the requirement that the decision maker is fully informed about the probability distributions before the decision at hand is already a high bar to clear, the paper makes an unsaid assumption that the decision maker has no predisposed biases irrelevant of data. This might explain why the first piece of data (shirt colors and gender) lined up so well with the model; no rational person has a built in bias for color of a shirt of gender (specially when pink is not an option). However, in something like the political dataset, the correlation between the model and the difference between groups might be impacted depending on the inherent biases of decision makers. This was not measured, as they were busy with providing evidence for the “kernel of truth” argument. In the current political climate, it is

unreasonable to believe that voters show up to a voting booth without already having some bias independent of data. That is, modern voters don't form stereotypes from representative types between groups, they form it some other way.

To critique the “kernel of truth” hypothesis, we focus in on the example of the stereotype that men are better than women at math. The example illustrates how the stereotypes comes from an unlikely difference of the right tail gender distribution of the SAT Math section. Thus, the model dictates that such a sharp difference in the right tail, however unrepresentative of the total population, will create exaggerated (and often incorrect) stereotypes. In this case, the stereotype produced is that men are better than women at math. However, the decision maker in this model knows every possible distribution for this trait, and thus he knows about different studies, some of which show no sharp right tail for men being better at math (such as when they are younger). Thus, how does the decision maker account for when two different studies or distributions give conflicting “kernels of truth?” Thus, I would argue that there are other factors, in this case a mix of misogyny and selective studies, that can impact how these stereotypes are formed. The other argument would be that these “kernels of truth” are simply products of the stereotypes. Thus, the stereotypes themselves create kernels of truth. In our example, this is equivalent to the idea that as girls grow up, they hear that they cannot be as good at math as boys, so they are discouraged and do not try as hard.

The rhetoric that won Donald Trump the election this year was thick with exaggerated stereotypes about immigrants. This model is able to explain where these stereotypes are founded in, and more importantly, to highlight the errors in these stereotypes. While the “kernel of truth” argument can be misrepresented as actual significant differences between Americans and Hispanics, the paper is clear that it is in a decision maker's perception of how sharp these differences are, no matter how unrepresentative these differences might be, that actually influence their decisions. Thus, this model provides reasoning to such exaggerated rhetoric while also making sense of some more representative stereotypes (such as tall Dutch people). Moreover, this paper explains how such stereotypes are based on context dependence, and thus the rhetoric that

highlighted the differences between distinct groups only served to fuel more exaggerated stereotypes. While this model might be overly simplified, it formalizes a subject in a way that not only works in some cases, but also sheds light on some properties of stereotypes which might not be obvious from today's rhetoric.

References

- [1] Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A., 2016. "Stereotypes." *The Quarterly Journal of Economics* 131(4).
- [2] Stats America https://www.statsamerica.org/sip/rank_list.aspx?rank_label=pop46&ct=S09