First impressions are not always correct
## Math helps measuring
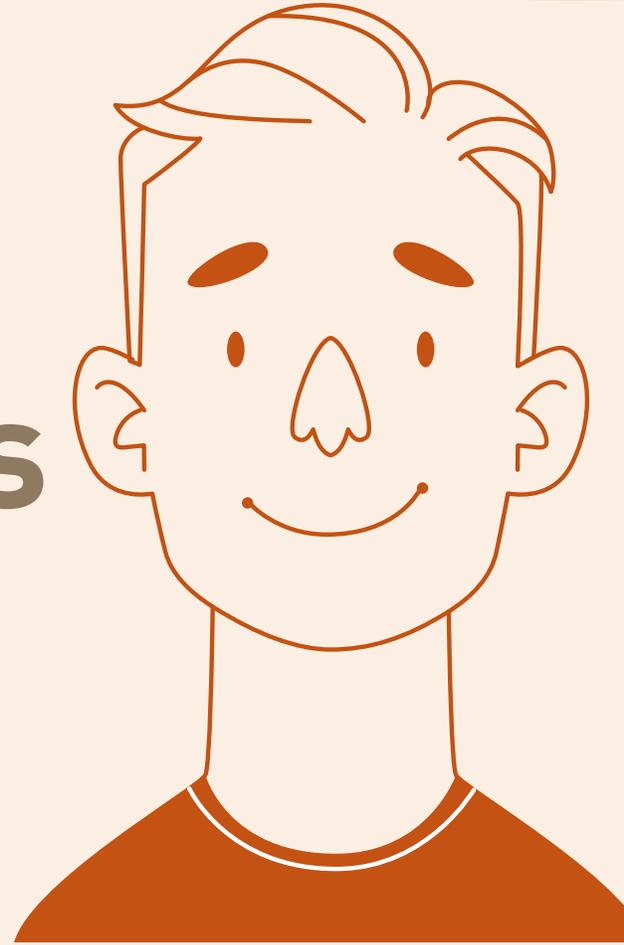
The use of mathematical models in psychology has become a key tool for the analysis of human behaviors. Sergio Barbosa, Professor of the Psychology Program at Universidad del Rosario, decided to identify a predictive model that best reflects human reasoning when judging moral attitudes.

By Camilo Calderón Acero
Photo Milagro Castro

In 1998, sitting U.S. President Bill Clinton was entangled in a political-sexual scandal when an out-of-wedlock affair with one of his White House interns was revealed. His favorable image among the citizenry changed radically and put his own term at risk. The interesting aspect about the case is that the public's attention was not focused on his performance in the position but on a moral aspect.

Many disciplines have been concerned with pondering and analyzing human behavior and how it should be based on ethics and morals, like in philosophy or sociology. Likewise, **moral psychology** has deepened its insights into understanding not only what we consider moral but also how human beings elaborate and vary (update) their judgment about what is right or wrong behavior.
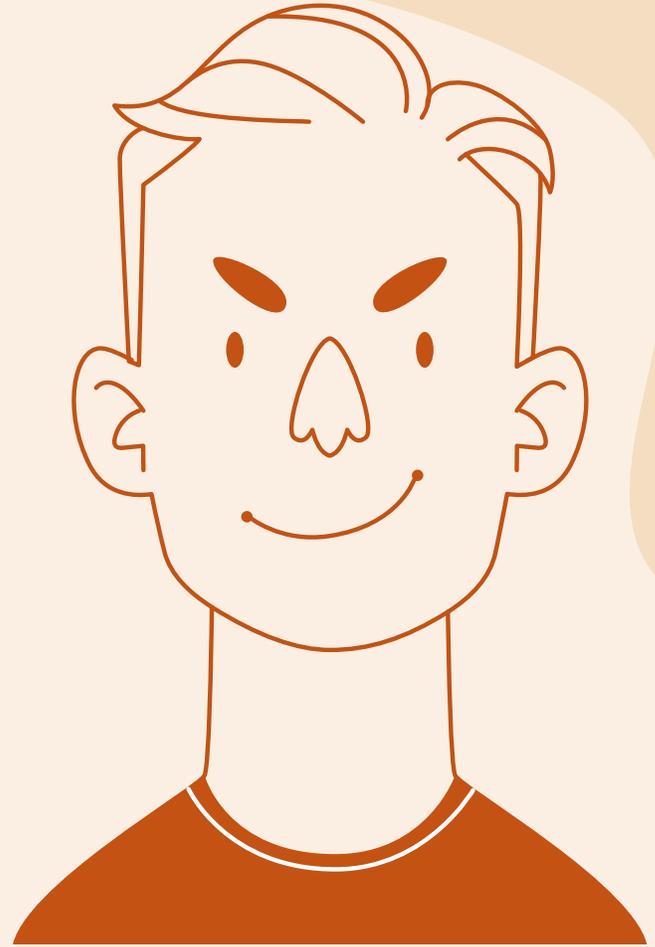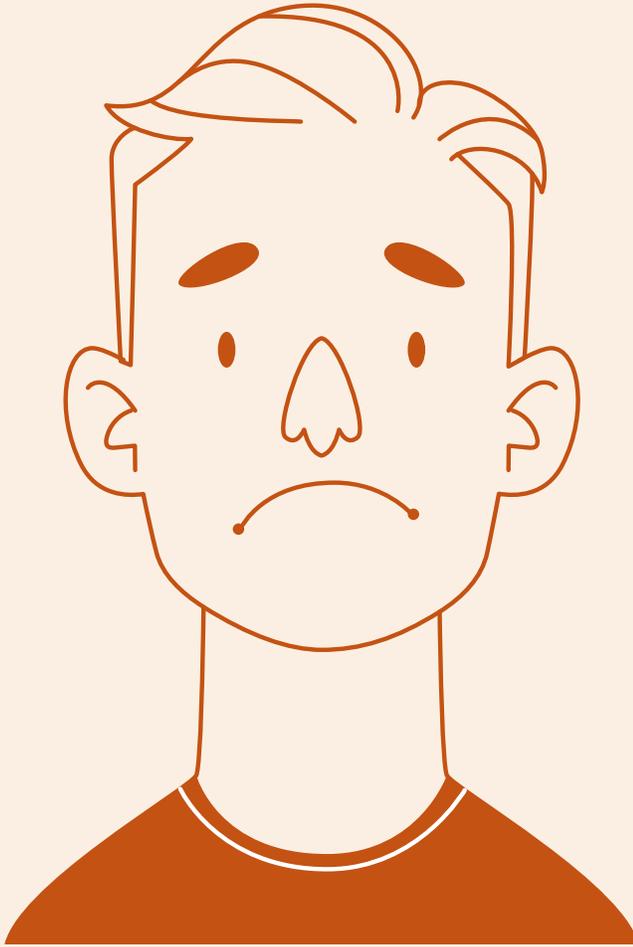
moral behaviors

To obtain a more accurate assessment of this phenomenon, Sergio Barbosa, Professor of the Psychology Program of the School of Medicine and Health Sciences of Universidad del Rosario, conducted research on the subject. He not only sought to delve into the way we attribute moral traits and how these traits are updated over time, but the research also showed that this phenomenon can be adequately captured by predictive models such as Bayes' theorem and similar models.

This allows to calculate the probability of occurrence of one event based on new information known from other related events.

For example, the probability of a rainy day may be 30 per cent, but that changes if we also consider that it is winter and that the area is being affected by the La Niña phenom-

↑
To be able to create a predictive model that could anticipate what people would say when faced with certain information or moral attributes, the first thing to do was to set up the parameters of both positive and negative attributes.

enon. Each new piece of information is likely to modify and update the model.

According to Barbosa, "as we give people more information, they should change their beliefs." That happened in our experiment: people did not stick to their first impression. The interesting thing was that some participants would update their beliefs following a Bayesian model, especially for positive attributes. But when we added negative attributes, i.e., behaviors they considered to be those of a bad person, the models were mathematically simpler, based on averages.

Averages are faster to process, but they leave out very different values. For example, a student may have four scores - 4, 4.5, 1 and 3.7 - and their average will be 3.3.

This is far from being the student's actual performance because averages are highly hypersensitive to extreme values, high or low. Only having one very low score (1) affects the final value. With moral traits, this would be equivalent to making generalizations about someone based only on certain behaviors.

The interest in involving the exact sciences in this type of social science research is linked to the use of more precise methodologies, the results of which can be more contrastable. "We can define how the psychological process to be studied operates in mathematical terms and confront it with what happens in real life. In this case, we use it for the attribution of traits related to morals, vices, and virtues of the human being," emphasizes Professor Barbosa.

In the field of psychology, Bayesian models have been used to explain memory and perception phenomena, language acquisition, word segmentation, and interpretation of complex language, among others. However, this is one of the first studies that use those hypotheses for the analysis of moral judgments.

From the point of view of [William Jiménez Leal](), associate Professor in the Department of Psychology at Universidad de los Andes and also director of Barbosa's Ph.D. program, which was the origin of the research, "the use of probabilistic models allows us to establish a dialogue with other disciplines, such as mathematics, economics, or computer science, something that would be unlikely otherwise. The kind of predictive modeling makes it much more feasible that professionals in these disciplines are receptive to our research studies," he claims.

How to measure a certain behavior

In order to create a predictive model that could anticipate what people would say when faced with certain information or moral attributes, the first thing to do was to set up the parameters of both positive and negative attributes.

For this purpose, they used open and closed questionnaires with psychology students; their answers allowed them to extract more than 100 morally relevant adjectives.

This sample was screened to eliminate synonyms, colombianisms, and offensive terms. Eight adjectives (four virtues and four flaws) arose from this screening process to finally reduce the adjectives most consistently classified as morally good or bad to four.

In the next stage, another group of students were surveyed to identify the most related behaviors; for example, those that would make them think of a person as dishonest. As in the first stage, more than a thousand terms were screened and with a final result of around 100, the final questionnaires were used with another group of people. At this point, a fictitious character, *Andrés*, was created to portray the classified behaviors.

The research study *"Different algorithmic models underlie virtue and vice attributions"* was approved for publication in Collabra magazine: Psychology journal of University of California, an open science Q1 journal.

"With all the information they had about *Andrés*, we asked respondents to rate from zero to one hundred how good or bad person they considered *Andrés* to be, and that is the data we compared with the predictive model," explains Barbosa. "What the model did was react to each of the situations we introduced to the respondents."

The parameters of the model were given by the definition of prior probability (how much of this trait do you think a person might have?) and conditional probability (how much do

↑
"We can define how the psychological process to be studied operates in mathematical terms and confront it with what happens in real life. In this case, we use it for the attribution of traits related to morals, vices, and virtues of the human being," emphasizes Professor Sergio Barbosa from the School of Medicine and Health Sciences of Universidad del Rosario.

you think a person displaying this behavior might have such a personality trait?). In the end, these model responses were cross-checked with the results of the student questionnaires.

The math of a moral judgment

The first comparative results were slightly different from expected. In his initial hypothesis, Barbosa had stated that the cognition process executed when signaling moral judgments would be updated following a neutral Bayes' Law or similar model, where people would not stick to their first impression of someone but would update it following Bayes' Law.

However, the researchers found that the type of model used would vary based on the information they received from *Andrés*. This apparently simple distinction became more complex when the respondent was faced with seemingly opposite behaviors. This incongruence between a person who at the same time could have two opposite attributes in different consecutive questions showed that the respondent often chose Bayesian models, the more complex ones.

According to Professor Barbosa, this happens because "when the information is more congruent, it is easier to be interpreted, since it all goes to the same side. When I have more

incongruent information, the more I use a Bayesian model . These models tend to be more accurate, but are more cognitively expensive. I would not strive to use these complex models when I have simple, easy-to-interpret information. I tend to use Bayesian models when the information is complex and I must make sense of seemingly incompatible elements."

For example: How can I state whether *Andrés is* a good or a bad person from my knowledge that he likes cooking, he spends a lot of money on gifts but also usually parks his car in handicapped spaces? What about if he also 'queue jumps' but treats the elderly with respect?

According to [Juan Samuel Santos](), professor of the Faculty of Philosophy at the Pontificia Universidad Javeriana, this sort of study is useful to the extent that it allows us to understand and be aware of mistakes when making moral judgments, especially when they are influenced by particular circumstances. "Flesh-and-blood people often make mistakes or do bad things, and applying a simple model to judge only certain behaviors can lead us to misjudge the actual person," he says.

It is worth mentioning that, although mathematical models of this style allow for establishing aggregate trends, as in the case of this research, it is natural that they have a considerable margin of error when analyzing the decisions of specific individuals.

Scopes and risks

Reflection on human behavior is becoming increasingly important in academia, given the impact of external elements that can determine changes in points of view. Every day, cases of disinformation and media slander based on a video, an image, or a text in social media are widespread. In the public debate, it is even quite common to hear the term cancellation culture, which is based primarily on the rapid dissemination of judgments on the behavior of others.

In scenarios like that, the use of predictive models can be a very useful tool to better understand how the aforementioned phenomena operate. As psychologist William Jiménez Leal points out, "we are exposed to constant manipulation by social media that shape moral opinions. As long as we understand which are the things that are to be manipulated, we can make truly autonomous decisions."

New challenges also arise from this, as a predictive model can be created that is useful in detecting the nerves to hit and how to provoke indignation or favoritism towards a person or social group.

According to Professor Santos, manipulation is an inevitable condition in the discovery of new knowledge, but it can also be counteracted. "If someone wanted to manipulate the image of a public person to damage them, they would show situations in which that person appears negatively more often. This would lead the audience to use simpler models and lead to conclusions about that individual's behavior that are quicker and more susceptible to error," he asserts.

Hence, further study of predictive models in moral psychology can be applied to updating beliefs in matters of social

Bayes' theorem and similar ones allow us to calculate the probability of an event occurrence based on the new information known from other related events. For example, the probability of a rainy day may be 30 percent, but it changes if we also consider that it is winter and that the area is going through the La Niña phenomenon. Each new piece of information can modify and update the model.

From his perspective, Santos urges us to consider that other research in the field of moral psychology has found limitations, because "human beings are rational, but bad calculators; and although they should update their beliefs according to theorems such as Bayes', the human mind does not actually work that way."

This perspective can also reduce errors when it comes to research and to further deepen their understanding.

This is why Professor Barbosa's work in the Psychology Program at Rosario will keep on studying how moral judgment updates work, especially for complex issues. And to achieve this, the joint work that he develops with his students within the Research Incubator of Moral Psychology and Decision Sciences is fundamental.

Within his initial interest in 2016, when he began this study, the effect of the peace process on Colombian society was absolutely present, and it is an issue that is still very current today. The challenge of reconciliation will continue to be a central issue in the country for the next few years, and to understand it, to better understand its development, the contributions by various social disciplines are essential. ◻

reconciliation, xenophobia, or any kind of discrimination, phenomena that in countries like Colombia are still strongly present and that tend to be exacerbated in times of strong social commotion.

"We would be interested in modeling how people distinguish between '*I was told*' and '*I witnessed it.*' '*I was told about it by a source I trust, a neutral source, or a source I don't trust.*' This is one of the most important variables in today's communication, where people tend to only listen to certain sources, those that are more reliable, and discard the others. We are interested in going deeper into how this works in mathematical terms," emphasizes Barbosa.

↑
New challenges arise as a predictive model can be created that is useful in detecting the nerves to hit and how to provoke indignation or favoritism towards a person or a social group.