

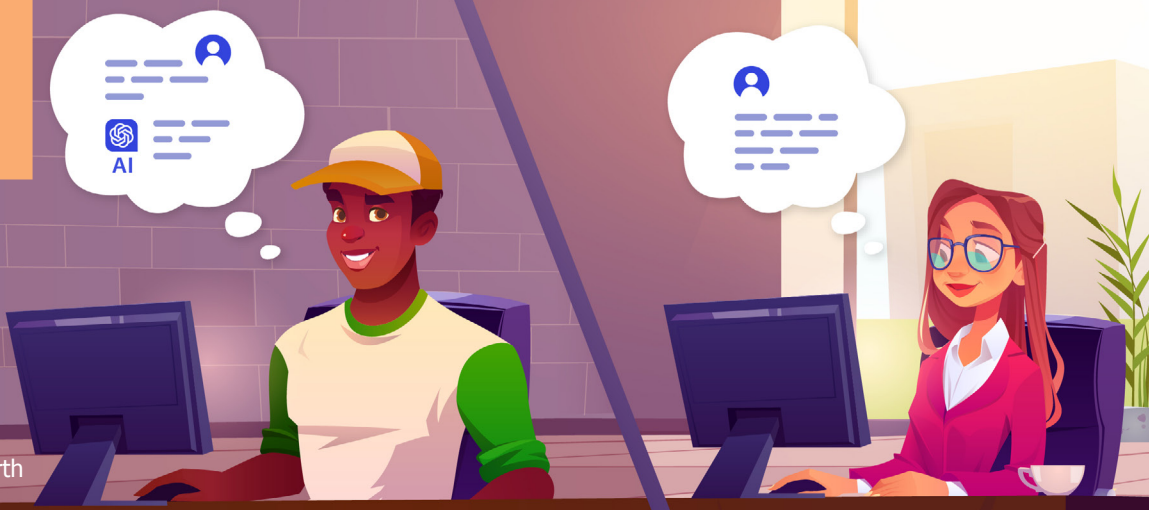
How can AI make online chats kinder?

Authors:

Lisa Argyle, Christopher Bail, Ethan Busby, and others

Associate Editors:

Elitsa Panayotova and Fiona Firth



Abstract

Imagine a world where everyone can talk and share their ideas. Even if they don't agree, they listen to each other. This is what should happen in a democracy. But sometimes, when people chat online, things can get a bit mean. We wanted to see if artificial intelligence (AI) could make online talks kinder.

So, we did an online experiment with pairs of people discussing a tricky topic: laws about who can own guns. One participant in each pair had an AI assistant that made

suggestions. The participants could accept or ignore the suggestions. Most participants took the AI's advice, which made people more positive and polite. Even if they disagreed, they understood and respected each other. The good thing is that the AI didn't change the topic, it just helped people to talk to each other more kindly.

Introduction

Have you ever wondered what **democracy** is all about? At its core, it means that everyone has a voice and can express their opinions, and the government pays attention to what people want. This also means that everyone should let others express their views, even if they don't agree. People **deliberate** – they try to find solutions to problems together. For example, imagine that you and your friends have to decide where to go for a fun day out. You would prefer that everyone listens to each other instead of arguing, right? In a democracy, we call talking and sharing ideas like this **political discourse**. Usually, these deliberations happen face-to-face – in a massive town hall meeting for example.

To keep everything polite between so many people, there is a **moderator**. Their job is to make sure things go smoothly. Nowadays, a lot of these discussions happen online. People send billions of messages every day on social platforms. These **online conversations can have a huge impact and shape how societies make decisions**. But there is a growing problem. Sometimes people don't listen to each other and can say mean things. **Wouldn't it be better if there was a moderator in online talks, too?** We tried to use **artificial intelligence (AI)** to carry out this task. Would it be helpful? Would people be nicer and the discussion better? This is what we wanted to find out.

Methods

We did an online chat experiment. We recruited a diverse group of participants and asked them to talk about **gun regulation**. This topic is very divisive and usually generates heated discussions.

The participants talked in pairs and only one of them had

an AI assistant. The AI suggested a different approach that helped people understand each other (Figure 1). It made these suggestions based on previous work on moderated discussions. We know that face to face discussions work better with a moderator. So, **just like a moderator, the AI**

tried to make sure the discussion was polite and respectful. It didn't change the participant's point of view, though! They could accept or refuse the suggestions or completely edit the message. We also had a **control group**. In this group, the people had a conversation on the same topic but neither of the participants had an AI assistant. This allowed us to

assess the AI's helpfulness.

After the discussions ended, the pairs shared their thoughts on how good the conversation was. For example, did they feel like they understood the other person? How did they feel about gun regulation afterward?

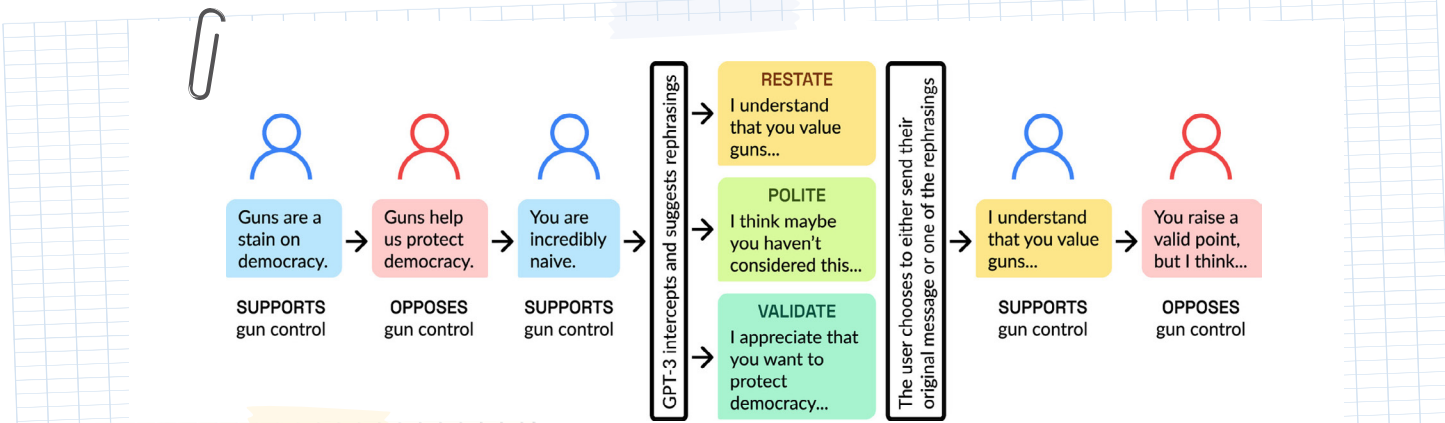


Figure 1:
An example of an online conversation about gun regulation. The AI chat assistant looks at the user's message. It suggests a kinder alternative without changing the topic or point of view.

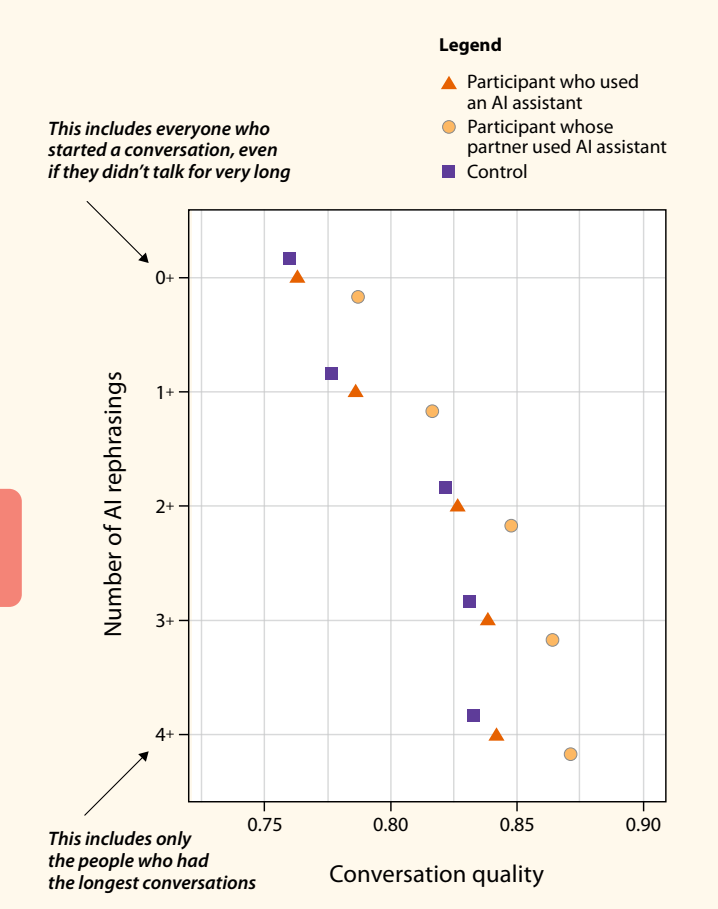
How did you feel when you read this exchange? Did you immediately take a position one way or another? If so, this could be due to **social polarization**.

Results

The participants accepted the AI's suggestions two thirds of the time. When they used the AI's suggestions, the messages became more positive and polite. Yet the main topic stayed the same. Participants whose partners used the AI reported feeling better about the conversation. They were more willing to understand and respect each other, even if they still disagreed. The people who used the AI themselves didn't report a huge change in how they felt about the conversation. However, they were more willing to respect the opinions of others! (See Figure 2.)

Which group reports the greatest conversation quality?

Figure 2:
The effect of rephrasing on conversation quality. The higher the number, the better the quality. Look closely at the Y-axis – the numbers increase going downwards.



Discussion

Online discussions can get a bit intense. Especially the ones about politics! The good news is our online chat assistant can be of great help. It acts as a digital peacemaker, nudging people to be more polite and respectful. It does all that without altering the main ideas that people discuss. It's almost like having a trained moderator.

In addition, the AI didn't force anyone to do anything. People could follow the suggestions or not. But here's where it's even better: the AI can jump in and give suggestions before you send your message! If it was a face-to-face conversation, you would have already said something mean.

Many people are concerned that AI might spread wrong information. Or it might make arguments online more extreme. **The results of our study show that AI can actually be a good thing in online discussions. It worked for both the person using it and their partner.** But it worked best for the partner. This is because the AI's job was to make the other person feel more understood and respected.

Conclusion

Online discussions sometimes get heated and lack respect. In our study, the AI did a good job of making online conversations polite. But you don't need an assistant to be respectful. Even in the vast digital world, each conversation matters. Your choice to be kind and respectful can have a ripple effect. It can promote a culture of healthy discussion

and mutual understanding. This is also true for real-life conversations. Our study does highlight the benefits of AI assistance. But it also emphasizes the importance of individual choices in promoting positive conversations.

Glossary of Key Terms

Artificial intelligence (AI) - a smart helper made by people. It's a computer (or a robot) that can learn from information, make decisions, and even understand and talk with humans.

Control group - the group in an experiment that doesn't get any special treatment. This is the baseline or normal group that helps scientists see if the changes they made to the other group (or groups, called treatment groups) really made a difference.

Deliberate/deliberation - a discussion in which the participants have a shared goal of solving a problem.

Democracy - a system where everyone has a say in making decisions. It's about working together, listening to different opinions, and making choices that consider everyone's needs and ideas.

Gun regulation - rules about who can have guns and how they can be used. They try to keep everyone safe and make sure guns are used responsibly.

Moderator - the person in a meeting who keeps things on track. They make sure everyone has a chance to speak, prevent arguments, and help the discussion run smoothly.

Political discourse - a conversation where people talk about important things like rules, ideas, and problems in a country. It's how we share our thoughts and try to understand each other to make our community better.

Social polarization - society is divided into strongly contrasting groups.

Check your understanding

1 Why is political discourse vital for democracy?

2 Why is it helpful to have a moderator (or an AI assistant) when people talk in a group?

3 What are some other ways AI could be useful in online conversations?

4 What could be some downsides of AI for people?

5 Divide people in your class into those who worry about AI and those who don't. Elect a moderator and discuss the concerns about AI as a group.

REFERENCES

Lisa Argyle, Christopher Bail, Ethan Busby, Joshua Gubler, Thomas Howe, Christopher Rytting, Taylor Sorensen and David Wingate (2023) *Leveraging AI for democratic discourse: Chat interventions can improve online political conversations at scale*. Proceedings of the National Academy of Sciences.

<https://www.pnas.org/doi/10.1073/pnas.2311627120>

Kiddle: Artificial intelligence facts for kids

https://kids.kiddle.co/Artificial_intelligence

United Nations: Democracy

<https://www.un.org/en/global-issues/democracy>

Acknowledgment:

This article's adaptation was supported by the Akamai Foundation.

