Self-Persuasion: Evidence from Field Experiments at International Debating Competitions

Peter Schwardmann (Carnegie Mellon University)
Egon Tripodi (University of Essex and JILAEE)
Joël J. van der Weele (University of Amsterdam and Tinbergen Institute)

Discussion Paper No. 306

December 1, 2021

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Self-Persuasion: Evidence from Field Experiments at International Debating Competitions*

Peter Schwardmann Egon Tripodi Joël J. van der Weele

November 4, 2021

Abstract

Laboratory evidence shows that when people have to argue for a given position, they persuade themselves about the position's factual and moral superiority. Such self-persuasion limits the potential of communication to resolve conflict and reduce polarization. We test for this phenomenon in a field setting, at international debating competitions that randomly assign experienced and motivated debaters to argue one side of a topical motion. We find self-persuasion in factual beliefs and confidence in one's position. Effect sizes are smaller than in the laboratory, but robust to a one-hour exchange of arguments and a ten-fold increase in incentives for accuracy.

JEL Classification: C93, D72, D83, D91.

*Schwardmann: Carnegie Mellon University; schwardmann@cmu.edu. Tripodi: University of Essex and JILAEE; egon.tripodi@essex.ac.uk. van der Weele: University of Amsterdam and Tinbergen Institute; j.j.vanderweele@uva.nl. We are grateful to co-editor Stefano DellaVigna, a second co-editor and three anonymous referees for comments that greatly improved the paper. We also received valuable comments from numerous seminar participants and thank Saskia Bergmann, Andrea D'Souza, Yimin Ge and Lena Martinovic as well as many other members of the debating community for their collaboration. We especially benefited from advice and support from Huyen Nguyen. Giacomo Manferdini, Hannah Rehwinkel and Irene van Rooij provided excellent research assistance and we are grateful to numerous local field assistants for their support with the data collection. Research funding from the CRC TRR 190 Rationality and Competition, the Research Priority Area Behavioral Economics at the University of Amsterdam, the Dutch Science Foundation (NWO), the European University Institute and the Russell Sage Foundation is gratefully acknowledged. This study was approved by the Economics & Business Ethics Committee at the University of Amsterdam (ref. 20181017051000) and by the Ethics Sub Committee 1 at the University of Essex (ETH2021-1082), and was pre-registered on the AEA RCT Registry (AEARCTR-0003922). Survey templates and replication files are available at https://osf.io/u7ekr/. We are convinced that this paper contains no errors.

> It might be plausibly maintained that in almost every one of the leading controversies, past or present, in social philosophy, both sides were in the right in what they affirmed, though wrong in what they denied; and that if either could have been made to take the other's views in addition to its own, little more would have been needed to make its doctrine correct.
>
> *John Stuart Mill, An Essay on Samuel Taylor Coleridge*

When asked to defend a particular point of view, people shift their private opinions in order to align them with the new arguments. Decades of research in the experimental laboratory have demonstrated this effect by having subjects argue in a randomly selected role (Janis and King, 1954; O'Neill and Levings, 1979), make counter-attitudinal statements (Festinger and Carlsmith, 1959a; Elliot and Devine, 1994), advise others to buy inferior products (Chen and Gesche, 2017; Gneezy et al., 2020), or convince others of their own ability (Smith et al., 2017; Schwardmann and van der Weele, 2019; Solda et al., 2019). In experimental courtroom or bargaining settings where subjects argue a randomly selected side of a case, they adopt self-serving views of the underlying evidence that limit their willingness to compromise (Thompson and Loewenstein, 1992; Babcock et al., 1995; Engel and Glöckner, 2013).

This effect of persuasion goals on beliefs and attitudes, which we call self-persuasion, can have important implications. It limits the potential of communication in resolving costly disagreements in pre-trial legal bargaining and labor disputes (Babcock and Loewenstein, 1997) and helps explain why political polarization and partisanship are at record levels (Iyengar et al., 2019; Gentzkow, 2016), even though the internet has made it cheaper than ever to communicate with people that differ in background and ideology. It has also inspired theories about motivated cognition (Taber and Lodge, 2006; Bénabou and Tirole, 2016) and the social origins of reasoning (Von Hippel and Trivers, 2011b; Mercier, 2011).

Given the robust self-persuasion effect in the laboratory and its potential implications, an important question is whether the phenomenon carries over to more natural settings where it may be reduced by expertise and higher stakes or drowned out by contextual factors (List, 2003, 2006; Levitt and List, 2007). The key difficulty in the field is to disentangle the causality between private views and persuasion goals. We confront this identification challenge by conducting preregistered field experiments at international debating competitions. The competitions feature parliamentary-style debates on topical motions like the freedom of movement in the European Union, investment in geoengineering, and the regulation of big technology companies. Because debaters are randomly assigned to persuasion goals just before the debate, comparing the beliefs and attitudes of the two sides yields clean esti-

mates of self-persuasion. We survey debaters pre- and post-debate to measure three separate outcome variables: beliefs about motions-related facts, confidence in the relative strength of each debating position, and attitudes toward motion-related charities. Across two offline and two online competitions, 473 debaters from 58 countries filled in a total of 4854 surveys relating to 19 different motions.

Several features of the debating competitions make them particularly well suited for our purpose. The setting is natural to debaters who participate in several similar competitions each year, are skilled at the task of persuading and are highly motivated to be persuasive. Performance in these prestigious international tournaments is scored by experienced and impartial adjudicators and confers status within the debating community. These incentives resemble those of professionals in politics and law, and many famous politicians and lawyers honed their skills in competitive debating.[1] At the same time, the competitions allow us to maintain a high degree of control. We are able to incentivize our elicitations to assure that they reflect true beliefs and attitudes. We also precisely time our surveys to measure self-persuasion pre-debate, i.e. after debaters know their persuasion goals but right before the start of the debate, as well as post-debate, i.e. right after the hour-long exchange of arguments.

We find strong evidence for self-persuasion before the start of the debate. First, debaters are more likely to believe that a factual statement is true if the statement strengthens an argument supporting their position. Second, debaters become more confident about the relative strength of their debating position, as measured by the subjective probability that teams arguing the same side of the motion in other debates will win. In our third outcome measure, a monetary allocation task between motion relevant charities, we find only weak evidence for self-persuasion.

We provide two benchmarks for the size of these effects. First, we find that the size of our estimates from the field is about 21 percent of the average effect in the laboratory studies. This may have to do with the selection of particularly polarizing issues in lab studies or with publication bias in the literature (Andrews and Kasy, 2019; DellaVigna and Linos, 2020). Second, we contrast the polarizing effect of self-persuasion with that of political partisanship, two effects that coincide in most settings but can be separated in this context. We find that the polarization predicted by the political attitudes of debaters is smaller than the self-persuasion gap on two out of three outcome measures.

Next, we investigate the impact of the debate itself on the polarization induced by self-persuasion.

[1]For instance, prominent Brexiteers Boris Johnson and Michael Gove were president of the Oxford Union, a renowned debate club. Other prominent politicians who were part of debating societies include Nancy Pelosi, Jimmy Carter, Margaret Thatcher and John Major. See either the site of the NSDA (speechanddebate.org/alumni) or worldcollegiatefriends.blogspot.com/p/famous-former-debaters.html for partial lists of famous former debaters.

Ex-ante, both an increase and a decrease in polarization are plausible outcomes. On the one hand, the debate gives both sides access to the same arguments, so impassionate reasoners should converge on the same conclusions. Previous research shows that explicit prompts to focus on the opposing side of the argument can lead to more balanced argumentation (Lord et al., 1984; Perkins, 2019). Moreover, Levy (2021) finds that online exposure to counter-attitudinal news reduces affective polarization. On the other hand, the act of debating may reinforce persuasion goals and lead to further self-persuasion. In line with this idea, exposure to opposing views was shown to harden pre-existing views and attitudes both in the laboratory (Lord et al., 1979; Taber and Lodge, 2006) and on Twitter (Bail et al., 2018).

Comparing pre- and post-debate outcomes, we do not find evidence for a decrease in polarization. Two of our three outcomes, i.e. confidence in one's position and attitudes towards charities, show a slight increase in polarization post-debate, while factual beliefs show a slight decrease. However, none of the effects are sizable, resulting in very similar self-persuasion effects post and pre-debate, although we can not rule out modest in- or decreases. This is not because debaters are ignoring opposing arguments: We find that they report a higher tally of arguments for the other side than before the debate. We also find evidence for convergence among the partisan polarization of debaters, a dimension of disagreement that is not reinforced during the debate.

We investigate a number of additional aspects of self-persuasion. First, we find that more experienced and more successful debaters self-persuade less in factual beliefs, but not in confidence. Thus, while experience decreases the bias on some dimensions, it does not eradicate it. Second, we vary the cost of self-persuasion by implementing a ten-fold increase in the incentives for accuracy on the belief elicitations. We do not find evidence that higher incentives impact self-persuasion in factual beliefs, and while they reduce self-persuasion in the confidence in one's position, this effect is not statistically significant. Third, we find that expert adjudicators predict self-persuasion, generating an intriguing contrast between the persistence of the bias and the awareness of it in the community. The predictions also capture some of the substantial heterogeneity in the effect across motions, survey questions and outcome variables, suggesting that contextual information may be used to anticipate polarization (DellaVigna and Pope, 2018). Finally, we provide some insights into the "how" and "why" of self-persuasion. We find that self-persuasion is partly driven by a biased investment in arguments for one's own side, suggesting the phenomenon results from a failure to account for this biased investment (Thompson and Loewenstein, 1992). We also find that debaters whose beliefs

happen to be more aligned with their randomly assigned persuasion goal receive higher scores in the debate, in line with the existence of instrumental benefits of self-persuasion.

Our field experiments add to a small set of papers that study motivated reasoning in natural settings, which emphasize motives for belief distortion other than persuasion goals. Di Tella et al. (2007) show that the quasi-random assignment of property rights to squatters results in heightened pro-market beliefs, which is suggestive of motivated reasoning. Oster et al. (2013) and Ganguly and Tasoff (2016) show that some people avoid getting tested for serious diseases, and link their findings to models of self-deception in the service of reduced anxiety. Finally, Huffman et al. (2019) show that managers distort their memories of past performance feedback to maintain overconfident beliefs.

Our results show that self-persuasion is a highly robust phenomenon outside of the laboratory: It occurs among highly motivated subjects with years of debating experience and is not eliminated by an hour-long exchange of arguments or a ten-fold increase in incentives for accuracy. While the field effects are smaller than effects in the lab, we still find a sizable effect of about a quarter of a standard deviation for two of our three outcome variables. These findings suggest that self-persuasion is a significant and resilient contributor to polarization and disagreement on policy issues.

I Experimental setting

Competitive debating has a long tradition as a platform for civil discussion on important and controversial topics. The format is based on parliamentary practices and features the random assignment of debaters to positions on a given issue. Therefore, in contrast to debates between experts or politicians, competitive debates require participants to take a stance that may not correspond to their original views. Today, many universities have debating societies that organize local or international tournaments, the most prestigious of which include the North American, European and World Championships.

We conducted field experiments at four international debating competitions. The *Munich Research Open* and the *Erasmus Rotterdam Open* took place in the spring of 2019. We then collaborated with the *Amsterdam Open* in October 2020 and the *London School of Economics Open* in February 2021 for a second wave of data collection. Due to the Covid-19 pandemic, these last two tournaments took place online. Like most international tournaments, all four competitions follow the procedures of British Parliamentary (BP) debating. Debates feature two teams of two debaters each in the Proposition, who argue in favor of a given motion, and two teams of two debaters each in the Opposition,

who argue against the motion. Debaters are randomly assigned either to the Proposition or the Opposition of a debate and to a speaking order.[2] They are not allowed to research the motion's topic and have only 15 minutes to prepare their speeches.

The motions are designed by "chief adjudicators", who tend to be members of the debating community that are highly recognized for having won or having been adjudicators of the final stages of continental and world championships. Chief adjudicators aim at designing motions that are balanced, with reasonable arguments on both sides, and that pertain to topical issues in politics, such as immigration, climate change and the regulation of new technology.

Debaters at our tournaments are predominantly undergraduate and graduate students that are members of debating societies, but also include former students that have entered professional careers of various kinds. Most debaters participate in regular meetings of their debating societies and travel to many tournaments each year. They tend to have strong analytical skills, an ability to think on their feet and a breadth of knowledge.[3] The most illustrious debaters at our tournaments have had successful runs at the European and World Championships.

Our four tournaments were organized yearly by university debating societies. The typical costs of organizing a tournament include the compensation of the adjudicators and the technical team, location rental and catering. These costs are usually covered through registration fees and external sponsorship. We were able to recruit these prestigious tournaments for our research by offering sponsorship that covered a significant share of the tournament's organization costs. Moreover, our survey payments to debaters helped attract a larger number of teams, by lowering the effective registration fee for debaters. In order to remain attractive to elite debaters, it was essential that our research design had an especially light touch and did not interfere with standard debating rules and procedures. The only significant departure from standard protocols we needed to negotiate was the administration of our crucial pre-debate survey after the debaters had prepared and just before they started speaking.

A Research design

We collected data in the five preliminary rounds of each competition, except in Rotterdam where we skipped the fifth round for logistic reasons. Each debater participates in all preliminary rounds, ex-

[2]Online Appendix Table A.1 describes the eight roles in a debate and the order in which debaters speak.

[3]Further discussion of the characteristics of debaters that take part in this format can be found on the website of the American Parliamentary Debate Association.

cept for rare case where someone feels unwell or particularly uncomfortable with a motion. Debaters answered three main surveys: a baseline survey at the beginning of the tournament, a pre-debate survey right after preparation time and right before the start of the debate, a post-debate survey right after the end of each debate but before adjudicators' ratings are announced. Pre-debate and post-debate surveys are collected in each of the preliminary rounds. The random allocation of persuasion goals allows us to identify self-persuasion by comparing the outcomes of the pre-debate survey between the two sides of the debate. We then measure the same outcomes post-debate to study how debates affect the polarization due to self-persuasion.

Outcome variables. The main outcomes collected in our surveys are the following.

- **Factual Beliefs:** We elicited the probabilistic beliefs in factual statements related to the motion. Factual statements were constructed such that, if they were true, one side of the debate would find them "convenient" in support of their arguments. To interpret this belief as a measure of factual belief alignment with the proposition, we keep the raw reported belief for facts that favor the proposition and compute the complement belief for facts that favor the opposition. Thus, higher values of the resulting outcome reflect a stronger alignment with the proposition.

- **Confidence in proposition:** We elicited the subjective probability that a majority of parallel debates in the round (excluding the debater's own debate) will be won by the proposition side of the debate. In excluding the debater's own debate, we elicit the confidence in the strength of the case for the proposition, rather than confidence in own ability to persuade. Higher values of this outcome thus capture the perceived advantage of the persuasion goal of the proposition, independent of speakers' confidence in their own ability.

- **Revealed Attitudes:** We asked debaters to allocate money between a "neutral" charity and a charity that was aligned with one side of the motion. Each charity was described to respondents in a short paragraph on the same survey sheet. Debaters choose their preferred allocation out of nine possible allocations, displayed in order from least favorable to the neutral charity to most favorable. To interpret this choice as a measure of attitudinal alignment with the proposition, we keep the raw order of the debater's choice when the motion-specific charity is aligned with the opposition and invert the order when the motion-specific charity is aligned with proposition. Higher values of the resulting outcome capture alignment with the proposition.

For concreteness, consider the following example of a motion and the associated factual statement, charity and confidence question.

Example of motion: This house regrets the EU's introduction of the freedom of movement.

Factual statement: More than 35% of UK citizens interviewed for the Eurobarometer in 2018 think that the Schengen Area has more disadvantages than advantages for the UK.

Charity: ACT4FreeMovement campaigns for freedom of movement with EU citizens. Its goal is to increase the capacity of EU citizens to effectively secure access to and knowledge of their rights, as well as build public awareness and political support for mobile citizen rights.

Confidence statement: Excluding the debate happening in this room, in at least half of the parallel debates of this round, one of the two teams on the Government side of this motion will rank 1st.

In addition to our three main outcomes, we elicited several other variables. The baseline survey collected background information of debaters, including experience with debating, past achievements, political orientation, and basic socio-demographics.[4] In the pre-debate and (online) post-debate surveys, we also ask debaters to report the number of arguments available for each side of the debate. Among these arguments, we ask them to indicate how many can be considered very strong.

Incentives for accuracy. We incentivized factual beliefs and confidence measures with a binarized quadratic scoring rule that paid in lottery tickets. Depending on their report $r \in [0, 100]$ and the objective binary answer $R \in \{0, 1\}$, subjects receive a lottery ticket that paid off a monetary prize of M with the following winning probability

$$w(r, R) = 1 - \left(R - \frac{r}{100}\right)^2.$$

Our general instructions in the baseline survey used both the mathematical equation, a simple quantitative example, and an intuitive explanation that truthful reporting optimizes the likelihood of winning the monetary prize (see Online Appendix D).[5]

---

[4]The baseline survey also included some incentivized factual knowledge "decoy" questions about topics not related to the motions. These questions served to obfuscate the elicitation of Factual Beliefs related to the motions and not give away the topics of the motions that were still secret at that point.

[5]In theory, this randomized quadratic scoring rule is incentive compatible for all risk preferences (Hossain and Okui, 2013; Schlag and Van der Weele, 2013). Whether this is actually the case in practice is a matter of ongoing debate. In the

At the offline tournaments the belief elicitation prize M was 30 euro. At the online tournaments, we varied incentives between a small prize of 5 euro and a large prize of 50 euro, randomized at the team-round level.[6] This variation in the accuracy bonus M allowed us to investigate whether a higher cost of self-persuasion reduces its prevalence. At the end of the debate, we randomly selected one report incentivized with price M to be paid out to subjects, i.e. one report in the offline tournament and two reports in the online tournament.

For the attitude elicitation, subjects allocated up to 10 euro between two different charities, where the budget constraint was concave in order to discourage extreme choices. One of the choices was randomly selected and the experimenters made the charitable payments for this choice on the subjects' behalf.

Survey overview. Table 1 summarizes the timing and collection of outcomes in each survey, highlighting slight differences in implementation between the offline and online tournaments. The baseline survey takes place on the first day of the tournament before the start of preliminary rounds. In each round, the pre-debate survey is collected between the end of preparation time and the start of the debate, and the post-debate survey is collected right after the debate. Only after the post-debate survey is over do debaters receive the ranking from adjudicators.

From all tournaments we also collected the "ballot", the official score sheet summarizing the deliberation of adjudicators. This includes two main performance measures: the ranking of teams and the debaters' individual speaking performance. We also collected an adjudicator survey. At the offline tournaments this survey asked adjudicators to provide their own independent rating of each debater's persuasiveness post-debate. At the online tournaments we instead surveyed adjudicators pre-debate and asked Factual belief questions, their predictions of the average response on each side of the debate, and their prediction of the average allocation of charitable donations for our Revealed Attitude questions on each side of the debate.[7] We incentivized adjudicators' responses at the online tournament using the same scoring rule as for debaters and randomly selected one question to be payoff relevant at the end of the tournament.

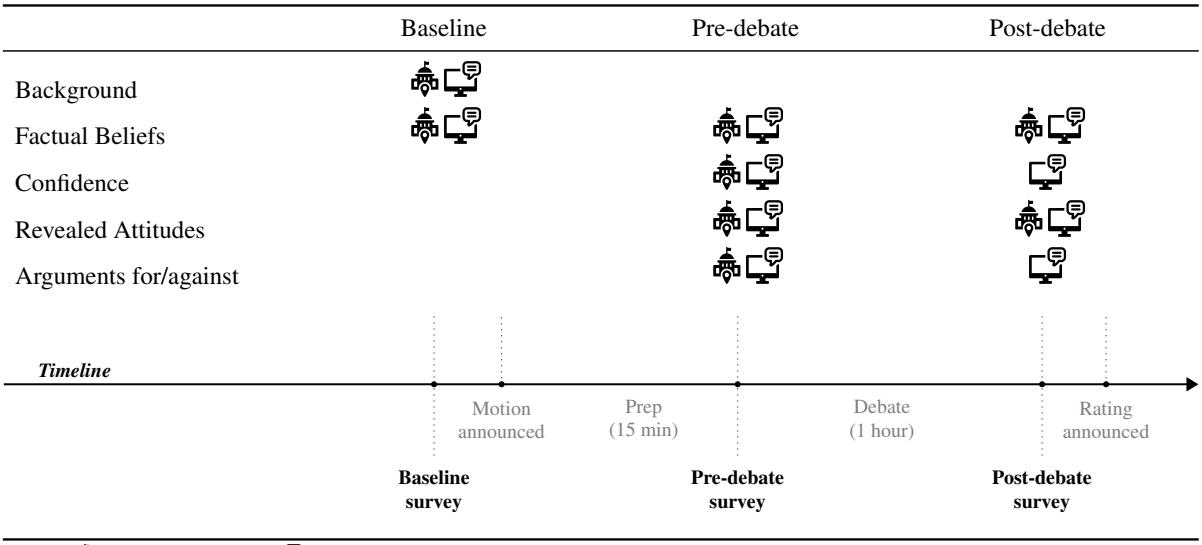
---

online tournaments, the formula and quantitative example were available upon clicking a box.

[6]The level of randomization was chosen in order to maximize the salience of differences in incentives. See Online Appendix D for further detail.

[7]The offline survey was designed to capture additional measures of persuasiveness. We only elicited this offline, as the post-debate survey interfered with adjudicators' deliberation about ballot scores and was difficult to administer online.

Table 1: Debater surveys: contents and timing



Notes: denotes offline and online tournaments.

The content of all surveys is described in greater detail in Online Appendix D, where Tables D.1 and D.2 provide all motions, factual statements and charities used for our elicitations.

Pre-registration. The first round of data collection was pre-registered on the AEA RCT registry (AEARCTR-0003922) with a pre-analysis plan. In a longer working paper, we execute the pre-analysis plan exactly and describe minor deviations (Schwardmann et al., 2019, pp 86-90). Additional hypotheses to be tested in the second wave of data collection, targeted sample size, and the alignment of factual questions and charities were also pre-registered as amendments to the plan.[8]

B Survey versions and administration procedures

Before the tournament, we coordinated with the chief adjudicators to converge on a final set of motions for the debate. For each motion, we developed several factual questions and motion-related charities, and varied the order in which factual questions and charities were presented to random subgroups of debaters.[9] The use of multiple questions in different orders means that debaters are not asked the same factual question twice. This helps rule out that results are driven by a desire to provide

[8] Amendments were submitted on November 27, 2020 (1 day before the Amsterdam Open), and February 5, 2021 (2 days before the LSE Open).

[9] See Online Appendix D for detailed description.

consistent answers to repeated questions and reduces concerns about experimenter demand effects. It also implies that no result depends on the answer to a single question or the order in which questions were asked. Moreover, since baseline and pre-debate questions were different both within and across subgroups, participants could not be influenced through discussion of the answers with others.

We administered the baseline after registration and introductory remarks by the organizers and research team, and shortly before the announcement of the first round motion. The full survey took about 25 minutes and was the same for all participants, except for the factual questions that related directly to the in-round motions, which were randomized. In offline tournaments, all debaters completed the survey in a single hall under the supervision of several enumerators ready to answer clarification questions. In the online tournaments, we administered the baseline in the virtual debate rooms where each enumerator was in charge of supervising eight debaters to maintain a high level of control and supervision.

In each debating round, the motions were announced in the central meeting room, after which debaters made their way to the assigned debating room. In the online tournaments, the central announcements took place via the app Discord, while debates occurred on Google Meets (Amsterdam Open) or Zoom (LSE Open). After the preparation period, enumerators distributed the pre-debate survey in the separate debating rooms. Debaters were given up to five minutes to answer the survey and enumerators ensured that they did not use this time to prepare for the debate. At the beginning of the debate enumerators also distributed the adjudicator survey which was collected along with pre-debate surveys at the online tournaments and after the debate at the offline tournaments.

After the pre-debate survey, the adjudicators opened the debate. The debate lasted about an hour and was attended by the enumerators. Once the adjudicators declared the end of the debate, enumerators distributed the post-debate survey, which was to be answered by debaters in five minutes.

C Sample characteristics and balance

On average, our sample has spent more than two years in debating, has qualified for more than seven quarter-finals of an international tournament, is about 21.5 years old, and tends to hold a relatively liberal ideology. The share of debaters that identify as women is 34.8 percent. The cultural background of participating debaters is fairly diverse: 61 percent are from Europe (including Russia and Turkey), 24 percent from Asia, 8 percent from North America, and 7 percent are either from Israel, Latin America, Africa, or Australia. Only 15 percent of participants are nationals of the country where

the tournament is hosted. In Online Appendix A.2 we show balance of individual characteristics and baseline alignment with the proposition across debaters with different persuasion goals.

II Main results

A Overview

We start with an overview of the dynamics of our main outcome variables (Factual Beliefs, Confidence, and Revealed Attitudes) across three points in time: at baseline, pre-debate and post-debate. The graph in Figure 1 displays the mean and 95 percent confidence intervals of each outcome for both proposition and opposition debaters. Histograms for the distribution of alignment for our three outcome variables pre and post-debate are provided in Online Appendix Figure A.1.

The first panel of Figure 1 shows the dynamics of Factual Beliefs, i.e. debaters' subjective probability that a state that favors the proposition is true. More information on which answers to factual statements, or states, are favorable to the proposition is provided in Online Appendix Table D.1. In the baseline survey, before debaters know the motion or before they are assigned to a side of the debate, Factual Beliefs of proposition and opposition debaters are identical. This implies that the randomization was successful. The pre-debate survey, taken after debaters prepared their arguments for 15 minutes, delivers evidence for self-persuasion. A 7 percentage point gap in Factual Beliefs opens up between proposition and opposition debaters. Self-persuasion persists in the post-debate survey, although the gap narrows to about 5 percentage points.

The second panel of Figure 1 displays the dynamics of debaters' Confidence in the strength of the proposition side of the debate. We see a clear gap of about 6 percentage points pre-debate, which widens to about 8 percentage points post-debate.

The third panel shows the results for Revealed Attitudes, measured by how much money the debater allocates to the charitable cause that is more aligned with the proposition. Recall that allocations were made along a concave budget constraint in nine discrete steps, so we use these steps as measurement units. We find a small pre-debate gap of about 5 percent of a discrete donation step. This gap then increases in the post-debate survey to about a quarter of a donation step.

Overall, Figure 1 shows clear evidence for self-persuasion both pre- and post-debate. The figure also hints at an unanticipated pattern: Average pre-debate and post-debate outcomes are tilted toward the proposition. This does not affect the interpretation of the main results because identification

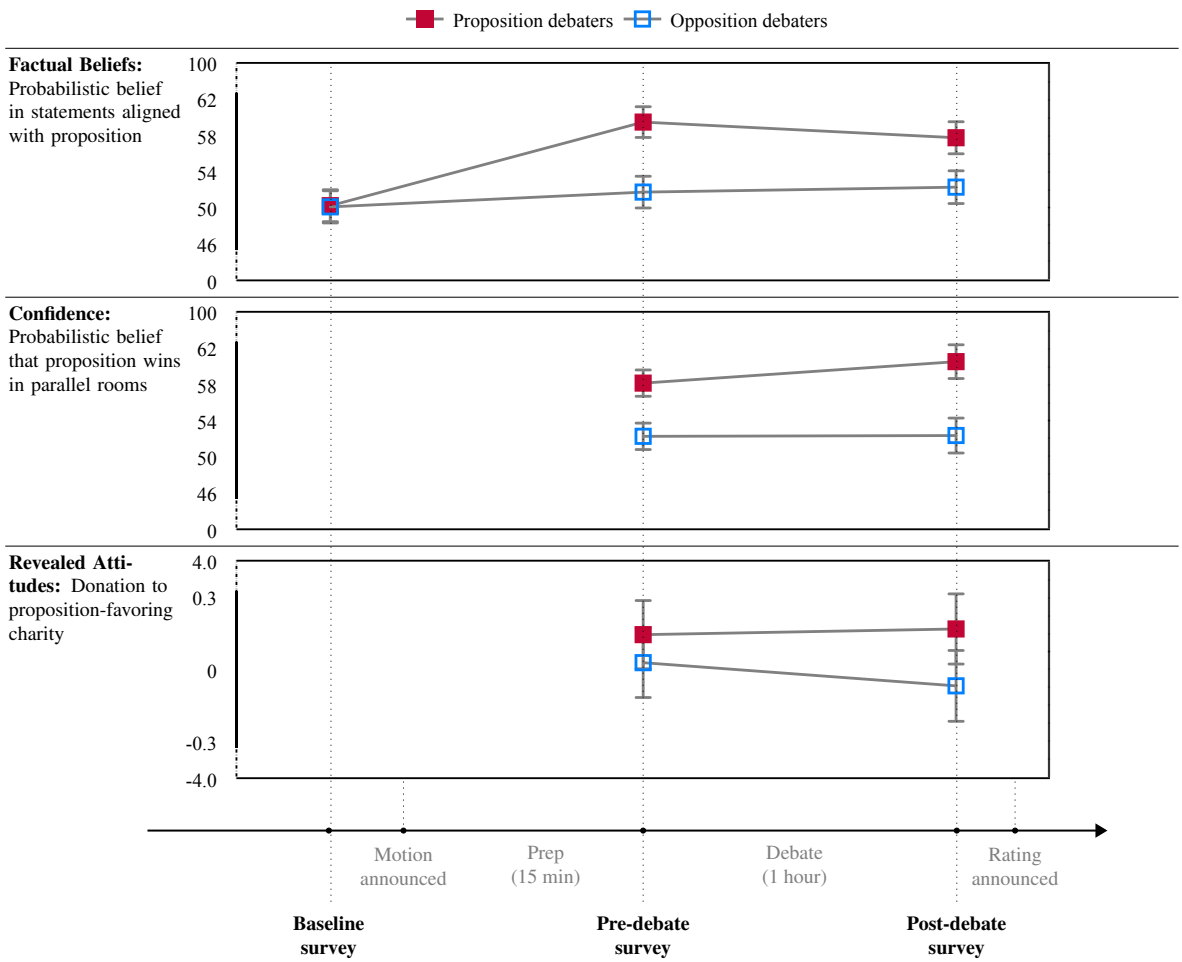


Figure 1: The dynamics of Factual Beliefs, Confidence and Revealed Attitudes.

Notes: The figure shows the dynamics of average alignment with the proposition for both sides of the debate, going from baseline, to pre-debate, to post-debate on our three main outcomes. For all three outcomes, higher values denote greater alignment with the proposition. The support of Factual beliefs and Confidence includes integers between 0 and 100, while Revealed Attitudes includes integers between -4 and 4. Each squared dot corresponds to the average of the alignment outcome at each point in time on each side of the debate and is placed between 95% confidence intervals. Dotted on the y-axis are segments of the support that are not plotted in the chart.

relies on between-subject comparisons of proposition alignment, conditional on the exact question or motion. However, one may wonder whether the asymmetry is due to a stronger self-persuasion effect among proposition debaters, or is driven by the characteristics of the motions or questions. To disentangle this, we elicit Factual Beliefs of a "control group" – the adjudicators of the online debate who knew the motions and answered the same questions as the debaters, but were not invested in the outcome. Using their answers as a benchmark, we find that proposition and opposition debaters self-persuade to an equal extent.[10] Furthermore, we find that the correct answer is aligned with the

[10]In the adjudicator's survey that we conducted in the online tournaments, we ask adjudicators to predict debaters'

proposition 47.3 percent of the time at pre-debate and 47.7 percent of the time at post-debate, so there is no strong imbalance. These findings suggest that debaters' beliefs simply happen to be more likely to favor the proposition.

B Pre-debate self-persuasion

We now turn to the statistical analysis of self-persuasion in the pre-debate elicitations. These effects reflect the cognitive processes taking place in the 15 minute preparation period after persuasion goals are assigned, but before the debate begins. We estimate self-persuasion effects in the following regression model

$$y_{i,q} = \alpha + \beta \mathbb{1}(proposition_{i,q}) + \delta_q + d_i + \varepsilon_{i,q} \tag{II.1}$$

where y is the outcome variable of interest with value $y_{i,q}$ for debater i answering question q, α is a constant, $\mathbb{1}(proposition_{i,q})$ is an indicator variable for being assigned to the proposition, δ_q is a question fixed effect, d_i is a debater random effect that is assumed to be orthogonal to the randomly assigned $proposition_{i,q}$, and the error term $\varepsilon_{i,q}$ is clustered within each team of debaters.[11] For Factual Beliefs and Confidence y is a number between 0 and 100, while for Revealed Attitudes y is a number between -4 and 4.

Table 2 shows in panel a the regressions for the pre-debate treatment effect, confirming the visual evidence. In addition, we present separate results for the two offline tournaments (panel b, columns 1-3) and the online tournaments (panel b, columns 4-6). These data differ in the time of collection, as the offline tournaments preceded the Covid-19 pandemic and the online data were collected during it, and in some details of the debating and survey procedures. All effects are robust to the omission of question fixed effects (see Online Appendix Table A.4).

In panel a, column (1) shows that proposition debaters are significantly more likely to believe that factual statements favoring the proposition are true and that statements favoring the opposition are false. Column (2) shows that proposition debaters are also significantly more confident that a majority of proposition teams will win the debates in the parallel rooms. The effects of Factual Beliefs and

---

beliefs and to state their own factual beliefs. Average Factual Beliefs of adjudicators are 54.6, which means they are slightly biased toward the proposition. These Factual Beliefs lie in the middle of proposition debaters' and opposition debaters' Factual Beliefs of 51.5 and 60.5 respectively.

[11] We cluster the error term at the team level because this is the level at which all randomization takes place. However, the error term might be also correlated at the room level, especially post-debate. Our main results are robust when clustering at the room level (see Online Appendix Table A.3).

Table 2: Pre-debate self-persuasion

Panel a): Full sample	All tournaments		
	(1)	(2)	(3)
Proposition alignment in:	Factual Beliefs	Confidence	Revealed Attitudes
Assigned to proposition	7.153	5.920	0.097
	(1.058)	(0.974)	(0.097)
Debaters	473	473	473
Observations	2217	2213	2212
R^2	0.216	0.110	0.194

Panel b): Offline vs. Online	Munich and Rotterdam (Offline)			Amsterdam and LSE (Online)		
	(1)	(2)	(3)	(4)	(5)	(6)
Proposition alignment in:	Factual Beliefs	Confidence	Revealed Attitudes	Factual Beliefs	Confidence	Revealed Attitudes
Assigned to proposition	6.192	4.389	0.277	7.821	6.943	-0.020
	(1.802)	(1.492)	(0.140)	(1.286)	(1.282)	(0.130)
Debaters	196	196	196	277	277	277
Observations	884	883	883	1333	1330	1329
R^2	0.140	0.034	0.136	0.268	0.157	0.223

Notes: Random effects linear regression model with standard errors (in parentheses) clustered at the team level. All specifications include question fixed effects. Each round, debaters are randomly assigned to argue either as proposition or opposition. The outcome is our measure of pre-debate alignment with the proposition in either Factual Beliefs, Confidence, or Revealed Attitudes. For all three outcomes, higher values denote greater alignment with the proposition. The support of Factual beliefs and Confidence includes integers between 0 and 100, while Revealed Attitudes includes integers between -4 and 4. The number of observations is determined by valid responses from debaters over 5 (4 in Rotterdam) rounds of debate.

Confidence are robust across tournaments, but are about 50 percent higher in the online than in the offline format, although the difference is not statistically significant.

For Revealed Attitudes, we do not see a statistically significant self-persuasion effect (panel a, column 3). In this case, there is a large difference in the online and offline format, with the effect being positive (and statistically significant) offline and negative and not statistically significant online. This appears to be partly driven by a survey design issue in the online format. Since the attitude elicitation was both the last and the most complicated survey item, time pressure and limited supervision may have reduced the attention of debaters to this question.[12] This interpretation is in line with the fact that the effect is present offline, where supervision was stricter, and post-debate, where time pressure was lower.

[12]We see that the debaters online are much more likely to favor the neutral charity relative to the motion-related charity, where the latter happened to be aligned with the proposition slightly more often. In exploratory analyses we exclude 73 debaters who did not answer all surveys and hence might not have taken it seriously. In the equivalent of column 3, we find a substantially larger self-persuasion effect (0.138, $p = 0.187$).

Effect size. How should we think about the size of these self-persuasion effects? We provide two benchmarks to answer this question. First, we can compare the standardized effects to those of the rather sizable laboratory literature on this topic. Table 3 provides an overview of laboratory work on self-persuasion.[13] By comparing the standardized effect sizes in the final column, we see that our field estimates are smaller than those of most other studies, although they are larger than estimates in Solda et al. (2019) and comparable to Schwardmann and van der Weele (2019). The relatively smaller effect size in our field setting mirrors the results in DellaVigna and Linos (2020), who show that the effects of "nudge unit" interventions in the field are about 25 percent of those obtained in comparable academic studies. They attribute this to publication bias in the academic literature and differences in the details of the intervention. In addition, we conjecture that the laboratory studies may haven been designed and piloted to contain sufficient scope for self-persuasion. In Section III.C, we show that there is a large heterogeneity in self-persuasion between motions, making topic selection consequential.

A second way to benchmark effect sizes is to compare them to the degree of political polarization in the outcome variables. Political ideology typically coincides with persuasion goals in the field, but can be separated here due to the orthogonal assignment of persuasion goals. To investigate the degree of political polarization, we construct a dummy variable that takes a value of one if debater's political leanings are aligned with the political leanings of a motion's proposition. We elicited debaters' political leanings on a left-right scale from 0 (very left) to 10 (very right). We classify a debater as left-leaning if their reply falls into the range of 0 to 4, and as right-leaning otherwise. According to this classification 26.4 percent of debaters are right-leaning.

We classify a motion as left-leaning if and only if, at baseline, left-leaning debaters are more likely to believe in the factual statements that support the proposition. Basing our classification on the revealed viewpoints of our debaters at the start of the tournament has the advantage of incorporating the political perceptions of the people whose partisan attitudes we are investigating. However, all our results are robust to an alternative political classification of motions, based on the ratings of an independent sample of 23 debaters (see Online Appendix Table A.5 for a detailed overview).

[13] We focus on self-persuasion in situations where subjects are a) incentivized to persuade or negotiate with others, and b) face incentivized belief measurements, as Bullock et al. (2015) shows that accuracy incentives in surveys reduce polarization effects. While we are not aware of other studies that fit these criteria, we do not guarantee this is an exhaustive list. We also include Festinger and Carlsmith (1959b) as a seminal reference point, although we cannot compute standardized effect sizes as no details of the sample distributions are reported, an issue that also plagues other early studies in psychology, like O'Neill and Levings (1979).

Table 3: Literature review effect size

Paper	Context	Persuasion objective	Treatment	Control	Outcome	Sample	TE	TE/ σ_y
1. Festinger and Carlsmith (1959a)	Boring effort task	Convince others that the task is enjoyable	Incentive 20 USD	Incentive 0 USD	Self-reported interest in the task (likert scale -5 to 5); unincentivized	Laboratory subjects; N=40	0.4	n.a.
			Incentive 1 USD	Incentive 0 USD	Self-reported interest in the task (likert scale -5 to 5); unincentivized	Laboratory subjects; N=40	1.8	n.a.
2. Thompson and Loewenstein (1992)	Fictitious wage bargaining	Negotiate favorable settlement	Submit wage offer as Union	Submit wage offer as Manager	Fair wage (in USD); unincentivized	Laboratory subjects N=40	0.15	0.780
3. Loewenstein et al. (1993)	Fictitious trial	Negotiate favorable settlement	Argue Prosecutor side	Argue Defendant side	Fair settlement (in USD); unincentivized	Laboratory subjects; N=160	17710	1.086
					Judge prediction (in USD); unincentivized	Laboratory subjects; N=160	14527	0.834
4. Babcock et al. (1995)	Fictitious trial	Negotiate favorable settlement	Don't know persuasion objective before reading materials	Know persuasion objective before reading material	Within pair difference in fair settlement (in USD); incentivized	Laboratory subjects; N=94	26031	1.087
					Within pair difference in prediction of judge (in USD); incentivized	Laboratory subjects; N=94	25491	0.932
5. Chen and Gesche (2017)	Financial advice game	Make advisee buy asset A over alternatives.	Commission to recommend Asset A	No commission to recommend asset A	Own choice of whether to buy asset A (binary); incentivized	Laboratory subjects; N=99	0.173	1.150
6. Gneezy et al. (2020)	Financial advice game	Make advisee buy asset A over alternative.	Review asset before learning commission for Asset A	Review asset after learning commission for Asset A	Belief that advisee prefers asset A to B (binary); unincentivized	Amazon Mturk workers; N=900	0.338	1.519
7. Schwardmann and van der Weele (2019)	Verbal persuasion task	Persuade others verbally of high test performance	Know about persuasion task	Does not know about persuasion task	Belief about own IQ test score (probability); incentivized	Laboratory subjects; N=688	0.060	0.309
8. Solda et al. (2019)	Written persuasion task	Persuade others in writing of high test performance	Already completed persuasion task	No awareness of persuasion task	Belief # correct answers (0-31); incentivized	Amazon Mturk workers; N=600	0.650	0.110
9. This paper	High profile debating competition	Win debate	Argue for motion	Argue against motion	Belief that facts in favor of the motion are true (probability); incentivized	Expert debaters; N=2217	0.078	0.264
					Confidence that teams arguing in favor of the motion win debates (probability); incentivized	Expert debaters; N=2213	0.058	0.227
					Donations towards charitable organizations supporting causes in favor of the motion (rank 0-9); incentivized	Expert debaters; N=2212	0.116	0.048

Notes: This table presents treatment effects (TE) and standardized treatment effects (TE/σ_y) from related experimental paradigms. The only study that we know has undergone exact replication is Babcock et al. (1995). The average standardized effect in the replication of Hippel and Hoeppner (2019) is about half the original effect (in the original study, effect sizes on outcomes 1 and 2 are 1.087 and 0.932, respectively; in replication effect sizes on outcomes 1 and 2 are 0.646 and 0.390, respectively), but still sizeable. The unweighted average of the standardized effect size for studies 2-8 is 0.867. The average effect size across the three main outcomes of the present study is 0.180, which is 20.8 percent of the unweighted average of the rest of the literature.

Table 4: Political polarization

Time of elicitation:	Pre-debate			Post-debate		
Proposition alignment in:	(1) Factual Beliefs	(2) Confidence	(3) Revealed Attitudes	(4) Factual Beliefs	(5) Confidence	(6) Revealed Attitudes
Politically aligned with proposition	4.606	-1.367	0.379	1.392	-0.700	0.081
	(1.435)	(1.043)	(0.115)	(1.616)	(1.173)	(0.121)
Debaters	463	463	463	462	271	462
Observations	2178	2174	2173	2141	1277	2139
R^2	0.201	0.087	0.207	0.236	0.121	0.226

Notes: Random effects linear regression model with standard errors (in parentheses) clustered at the team level. All specifications include question fixed effects. Each round, debaters are randomly assigned to argue either as proposition or opposition. The outcome in columns 1-3 is our measure of pre-debate alignment with the proposition in either Factual Beliefs, Confidence, or Revealed Attitudes. For all three outcomes, higher values denote greater alignment with the proposition. We call debaters right leaning if they report political views on the 0-10 political scale above 4. For each round, we regress baseline Factual Belief alignment on being a right leaning debater and categorize the motion of that round to be right leaning if the regression coefficient is positive (in Online Appendix Table A.5 we conduct a different categorization based on a small follow-up survey and find the results to be qualitatively robust). 'Politically aligned with proposition' equals 1 if both the motion and the debater are left/right leaning, and 0 otherwise. The outcome in columns 4-6 is the post-debate alignment analog. The number of observations is determined by valid responses from debaters over 5 (4 in Rotterdam) rounds of debate.

We say that a debater is politically aligned with the proposition if both debater and proposition are left-leaning or right-leaning. To illustrate, a right-leaning debater is clearly aligned with the proposition "This House would suspend trade union powers and significantly relax labour protection laws in times of economic crisis". We can now estimate political polarization with regressions that are analogous to those for self-persuasion in Equation (II.1); we just replace the indicator for *arguing for* the proposition side of the motion with an indicator for being *politically aligned with* the proposition side.

The results of this exercise are presented in the first three columns of Table 4. Comparing the effect sizes of this exercise with the self-persuasion effects in Table 2, we find that for Factual Beliefs, political polarization is about two thirds of the self-persuasion effect. Turning to Revealed Attitudes, where the effect of self-persuasion is small, we find that the political effect is almost 10 times as large as the self-persuasion effect across all tournaments and slightly larger than the self-persuasion in Revealed Attitudes we see in the offline tournaments. Finally, we do not see partisan polarization on Confidence, where the point estimate is negative and statistically insignificant.

Thus, in two out of three outcome variables, we observe a larger effect of self-persuasion than of political polarization, indicating that the self-persuasion effect is a quantitatively important driver of polarization in this setting. Note, however, that political polarization in our setting may be less pronounced than in the general population, as ideological heterogeneity is relatively small, and not all motions evoke clear ideological difference between left and right.

Table 5: Post-debate self-persuasion and convergence

	Post-debate			Diff-in-diff		
Proposition alignment in:	(1) Factual Beliefs	(2) Confidence	(3) Revealed Attitudes	(4) Factual Beliefs	(5) Confidence	(6) Revealed Attitudes
Assigned to proposition	5.055	7.940	0.200	7.139	6.922	0.077
	(1.264)	(1.295)	(0.095)	(1.173)	(1.261)	(0.103)
Post-debate				0.455	-2.181	-0.093
				(1.418)	(0.792)	(0.093)
Assigned to proposition×Post-debate				-2.245	1.098	0.132
				(1.991)	(1.203)	(0.132)
Debaters	470	274	470	473	277	473
Observations	2171	1286	2169	4388	2616	4381
R^2	0.236	0.159	0.224	0.098	0.151	0.152

Notes: Random effects linear regression model with standard errors (in parentheses) clustered at the team level. All specifications include question fixed effects. Each round, debaters are randomly assigned to argue either as proposition or opposition. The outcome of columns 1-3 is our measure of post-debate alignment with the proposition in either Factual Beliefs, Confidence, or Revealed Attitudes. The outcome of columns 4-6 is our measure of alignment with the proposition in either Factual Beliefs, Confidence, or Revealed Attitudes–either at pre-debate or post-debate. For all outcomes, higher values denote greater alignment with the proposition. The number of observations is determined by valid responses from debaters over 5 (4 in Rotterdam) rounds of debate. Post-debate Confidence was collected only at online tournaments.

C Post-debate self-persuasion and convergence

We now turn to the post-debate survey. As we discussed in the introduction, there are two ex-ante plausible hypotheses about the dynamics of self-persuasion over the course of the debate. On the one hand, the pooling of arguments from both sides should lead impassionate reasoners to reach similar conclusions and reduce polarization. On the other hand, the very act of debating may reinforce the effect of persuasion goals and increase polarization at the end of the debate.

The visual evidence in Figure 1 suggests that the answer lies in the middle, with little overall evidence of either convergence or divergence. The statistical results are presented in Table 5. Columns 1-3 show the treatment effects of our three outcome variables post-debate for all tournaments combined, analogous to the first three columns of Table 2. We now see a sizable and statistically significant self-persuasion effect for all three variables. This includes the coefficient for Revealed Attitudes, which was not statistically significant pre-debate.

To estimate the size of convergence or divergence, the regression models reported in Columns 4-6 include a dummy for being in the proposition, a dummy for the post-debate survey, and the interaction between the two. The coefficients for the latter term give the size of the difference-in-difference between the pre- and post-debate treatment effects. We do not find evidence of consistent

or statistically significant convergence or divergence across the three outcome variables.[14]

Information transmission during the debates. Given the lack of movement of our main variables, one may wonder if the debates resulted in any information transmission at all. To investigate this question, we use some additional measurements. First, we asked each debater for the number of separate arguments for either side of the motion that they could think of.[15] We find that the total number of arguments cited by debaters increases by 20 percent over the course of the debate, from an average of 6.4 at pre-debate to an average of 7.7 post-debate, a difference that is statistically significant (t-test, $p < 0.001$). These numbers are not due to debaters exclusively generating more arguments for their side: the number of arguments for the *opposite* side increases over the course of the debate from 2.8 to 3.5, or 25 percent (t-test, $p < 0.001$). This shows that debaters learned new arguments during the debate. In Section IV.A, we investigate the role of the number of arguments in the development of self-persuasion in more detail. One might be concerned that post-debate self-persuasion stems from a desire to not "admit defeat" in front of the experimenter. Inconsistent with that interpretation, we observe that the reported share of arguments in favor of their own position shifts from 57.0 pre-debate to a more balanced 54.5 post-debate (t-test, $p < 0.001$).

Another way to investigate the impact of the debates is to look at their effect on political polarization. To this end, we compare the degree of political polarization pre-debate and post-debate. As can be seen in columns 4 and 6 of Table 4, post-debate Factual Beliefs and Revealed Attitudes reflect less political polarization than their pre-debate counterparts. Thus, in contrast to the polarization induced by self-persuasion, debates did have a clear mitigating effect on political polarization. Note that in contrast to the randomized persuasion goals, the dimension of political partisanship was not reinforced during the debate, where many subjects argued against their own political leanings. This may explain why political polarization, but not self-persuasion, declines during the debate, and suggests that persistence of disagreement requires the reinforcement of persuasion goals during the debate.

Persistence of post-debate effects. How long does the effect of self-persuasion persist? Our ability to answer this question is limited by the two days' length of our tournaments and our inability to

[14]Our study has 79 percent power to detect post-debate convergence of the size of 50 percent of the pre-debate gap in our pre-registered primary outcome, i.e. Factual Beliefs.

[15]At offline tournaments, we only measure this pre-debate. At online tournaments, we measure this both pre- and post-debate. Here, we only report the numbers for the online tournament, so we can focus on changes.

contact debaters afterwards. Nevertheless, we addressed this point in the final survey of the online tournaments, where we asked debaters again about their factual beliefs related to all five motions in the qualifying rounds. The factual questions were the exact same that we asked them at baseline, allowing us to see if factual beliefs shifted between the beginning of the tournament and the end of the qualifying rounds. This is a strong test of persistence, since a concern to be consistent may reduce the difference between the two elicitations.

In Online Table A.6, we investigate the self-persuasion effect on day two – for Factual Beliefs – with the same regression model as our main analysis. We find that this effect remains sizeable at 80 percent the post-debate effect size and statistically significant ($p = 0.003$). We show that this effect is not driven by the fifth round of debate, which took place on day two of each competition right before the final survey. Thus, we conclude that the self-persuasion effect persists until the next day, despite the intervening engagement in at least one unrelated debate.

III Heterogeneity

In this section we look at heterogeneity of the pre-debate self-persuasion effect across debater experience, incentives for accuracy, and the topics of motions.

A Experience and past success

Is experience or past success associated with less self-persuasion? Experience may allow people to learn and reduce behavioral biases, as has been documented in the case of the endowment effect (List, 2003). High-profile debate tournaments are uniquely suited to study this question: While all participants have some degree of experience, there is still substantial heterogeneity in the number of years debaters have been debating as well as in their past successes, measured by how many times they previously made it out of the preliminary rounds into to the semi-finals at big tournaments.

In Table 6, we present regressions for Factual Beliefs and Confidence, the two outcome variables where we find significant pre-debate self-persuasion. In columns 1 and 2, we interact the treatment (being assigned to the proposition) with an indicator for having more than the median years of debating experience. In columns 3 and 4, we interact the treatment with an indicator for more than median number of semi-final attainments. These two binary indicators capture related aspects of experience (correlation $\rho = 0.486$), and the interaction terms with the treatment show how experience correlates

Table 6: Heterogeneity and stakes

	Experience and achievements				Stakes of elicitation	
Proposition alignment in:	(1) Factual Beliefs	(2) Confidence	(3) Factual Beliefs	(4) Confidence	(5) Factual Beliefs	(6) Confidence
Assigned to proposition	9.968	5.791	9.612	5.369	7.538	9.044
	(1.484)	(1.258)	(1.520)	(1.241)	(2.106)	(1.729)
Experienced	3.569	0.242				
	(1.698)	(1.793)				
Assigned to proposition×Experienced	-5.912	0.445				
	(2.109)	(1.994)				
High Achiever			2.587	0.974		
			(1.617)	(1.790)		
Assigned to proposition×High Achiever			-5.368	1.406		
			(2.076)	(1.842)		
High Incentive					1.196	2.219
					(2.112)	(1.907)
Assigned to proposition×High Incentive					0.598	-4.085
					(3.041)	(2.589)
Debaters	465	465	463	463	277	277
Observations	2187	2183	2177	2173	1333	1330
R^2	0.218	0.110	0.227	0.110	0.268	0.160

Notes: Random effects regression model with standard errors (in parentheses) clustered at the team level. Each round, debaters are randomly assigned to argue either as proposition or opposition. The outcome is our measure of pre-debate alignment with the proposition in either Factual Beliefs or Confidence. For both outcomes, higher values denote greater alignment with the proposition. In columns 1-2 we interact the treatment with Experienced, a binary indicator for above median years of experience in debating. In columns 3-4 we interact the treatment with High Achiever, a binary indicator for above median number of international tournaments in which the debater reached the kick-out phase. In columns 5-6 we interact the treatment with High Incentive, a binary indicator for randomly assigned incentive for the question. The number of observations is determined by valid responses from debaters over 5 (4 in Rotterdam) rounds of debate. In particular, remember that experimental variation in the stakes of elicitations was introduced only at online tournaments.

with self-persuasion.

In column 1, the experienced group shows about half the treatment effect on the Factual Beliefs measure, while for Confidence in column 2, we do not see an attenuating effect from experience. We obtain comparable results in column 3 and 4: High Achievers have about half the self-persuasion effect on Factual Beliefs, but there is not much difference in self-persuasion on Confidence and the sign of the effect is reversed. Note that the self-persuasion effect on Confidence for experienced debaters cannot be justified by their superior performance, as debaters were predicting the outcome of other, simultaneous debates, not their own. Thus, debaters of all levels of experience are subject to self-persuasion, although experienced and successful debaters show a smaller effect on the Factual Beliefs measure.

B Incentives

Economic theories of motivated cognition predict that beliefs are sensitive to the costs of misperceptions (Bénabou and Tirole, 2016). The previous literature has yielded some evidence for this prediction (Zimmermann, 2020), but there are also several null results (Mayraz, 2011; Coutts, 2019). In our setting, the cost of misperceptions is given by the monetary incentives for accuracy: the higher the deviation from the true answer, the lower the chance of winning the prize.

To test whether these incentives influence self-persuasion, we implemented exogenous variation in the prize that could be won. In the low incentive condition, debaters could win 5 euro with a correct answer, while in the high incentive condition, we increased the incentive 10 fold to 50 euros. These conditions were implemented within-subject and in the online tournaments only, with Factual Belief and Confidence questions randomly assigned to either condition. We informed debaters by displaying a "5 euro" or "50 euro" signal in front of the relevant questions.

Columns 5 and 6 of Table 6 show the regressions of pre-debate Factual Beliefs and Confidence on a dummy for being in the proposition, a dummy for being in the high incentive condition, and the interaction between the two. The coefficient for the interaction term shows that self-persuasion on Confidence is mitigated somewhat by high incentives, although the effect is not precisely estimated and not statistically significant. Similarly inconclusive results obtain from analogous regressions for the post-debate outcomes. Thus, we find no clear evidence for an effect of incentives.

A related question is how costly self-persuasion is to the debaters, in terms of foregone earnings from accurate answers. To this end, we compare the performance of the online debaters at pre-debate with a control group that did not have a persuasion goal. Like above, we use the adjudicators, who answered the same questions as the debaters in the online pre-debate survey. We find that adjudicators have a probability of winning the prize of 0.693, whereas debaters have a probability of winning the prize of 0.656 in the low stakes condition and 0.653 in the high stakes condition. This comparison implies that the cost of self-deception for debaters is 0.19 euro in the low stakes condition (with a 5 euro monetary prize) and 2 euro in the high stakes condition (with a 50 euro monetary prize).

C Heterogeneity across topics and the predictability of self-persuasion

The debates generate self-persuasion across multiple policy motions, which allows us to investigate the consistency of the effect. The three panels of Figure 2 show standardized self-persuasion effects

for Factual Beliefs, Confidence and Revealed Attitudes, aggregated at the motion level.[16] There are two main takeaways from this figure. First, there is a lot of heterogeneity in effect sizes across motions for a given outcome variable.[17] For example, debaters arguing for and against engaging private military companies to combat terrorism exhibited self-persuasion on Factual Beliefs that was more than three times larger than the average effect. Second, self-persuasion on one outcome does generally not predict self-persuasion on other outcomes.[18] In fact, it is hard to find a single motion for which we see (sizable) self-persuasion on all outcomes. Together these results suggest that studies should be careful in generalizing findings of polarization from any single topic or outcome.

The variability in the treatment effect across outcome variables may be explained by the variation in questions and charities selected by the researchers. However, we also see substantial heterogeneity in the treatment effect on confidence, where the elicitation question is always the same. This suggests the presence of a second source of heterogeneity that is related to the topic of the motions. This finding is in line with Tappin (2020), who identifies large variation in partisan political polarization across issues in U.S. politics.

Predicting motion effects. Is polarization predictable across motions and questions? The answer to this question will help us understand where and when disagreement arises and can potentially be avoided. As a first step towards answering this question, we investigate whether self-persuasion can be predicted by a group of experts. Such predictions also provide an additional benchmark for our effect sizes and help understand the awareness of the effect in the debating community (DellaVigna and Pope, 2018; DellaVigna et al., 2020).

We asked the adjudicators in the online tournament to predict the treatment effect. Adjudicators are arguably the best placed group to predict the effect of self-persuasion and of motion variation. They

[16]Figure 2 shows 15 out of 19 motions due to some of the motions' sensitive nature. Our analysis is based on all 19 motions.

[17]A test for heterogeneous treatment effects (Cochran, 1954) rejects the hypothesis that the treatment effect is homogeneous across rounds for Factual Beliefs ($Q-test$, $p<0.001$), Confidence ($Q-test$, $p<0.043$) and Revealed Attitudes ($Q-test$, $p<0.098$). To interpret the heterogeneity in treatment effects across rounds we estimate the I^2 (Higgins and Thompson, 2002), which measures the proportion of total variation across rounds that is due to heterogeneity rather than within-round sampling error. This measure is 62.2 percent for Factual Beliefs, 38.8 percent for Confidence and 31.0 percent for Revealed Attitudes.

[18]We cannot reject the null hypothesis that the correlation of standardized effects by motion across outcomes is zero. It is $\rho=0.154$ ($p=0.293$) between Factual Beliefs and Revealed Attitudes, $\rho=0.036$ ($p=0.882$) between Factual Beliefs and Confidence, and $\rho=0.048$ ($p=0.844$) between Revealed Attitudes and Confidence.

Table 7: Adjudicator predictions vs. actual effect sizes (online only)

	Factual Beliefs	Confidence	Revealed Attitudes
Actual effect size	7.81	7.22	-0.02
Predicted effect size	10.48	7.74	0.35
Correlation actual-predicted (motion level)	0.36	0.43	-0.07
Correlation actual-predicted (question level)	0.34	0.43	-0.04
Motions	10	10	10
Questions	30	10	20

Notes: Adjudicators' predictions are only available for online tournaments. Actual effect size at the motion level is calculated as the average proposition alignment among proposition debaters minus average proposition alignment among opposition debaters at pre-debate. Predicted effect size at the motion level is calculated as the average proposition alignment predicted by adjudicators minus average proposition alignment predicted by adjudicators among opposition debaters.

are experts in this particular context, as they have intimate knowledge of the debating environment and are experienced at debating as well as evaluating other debaters. In a pre-debate survey for each motion, we provided adjudicators with the factual questions and attitude elicitations related to the motion, and asked them to estimate the average responses for both proposition and opposition debaters. We incentivized their answers with the same scoring rule we used for the debaters, with a potential prize of 15 euros.

Table 7 compares the predictions with the actual (pre-debate) effect sizes, and shows that adjudicators do reasonably well in predicting pre-debate self-persuasion in the online tournaments. They overestimate the effect sizes for Factual Beliefs by about 34 percent and are strikingly accurate for Confidence. For Revealed Attitudes the adjudicators overestimate the mark substantially, as there was no effect in the online competitions, although their estimates are close to the actual effect in the offline competitions. Table 7 also shows the correlations of the predictions and actual effects on the motion level. An overview of the predictions organized by motion is given in Online Appendix Figure A.2.

The individual questions are another source of heterogeneity. In particular, for each motion, we have three different questions to elicit Factual Beliefs, and two different charities to elicit Revealed Attitudes. Some of these questions or charities may be more salient or have a stronger connection to the core arguments in the debate, and hence generate more self-persuasion. In the bottom row of Table 7, we show the correlations of the adjudicators' predictions and actual self-persuasion on the question level. Adjudicators have similar performance as for motions, showing that at least for Factual Beliefs, they are able to predict self-persuasion to some degree. Overall, adjudicators do a reasonable

job at not just predicting the existence of an overall effect, but at predicting its heterogeneity over motions and questions.

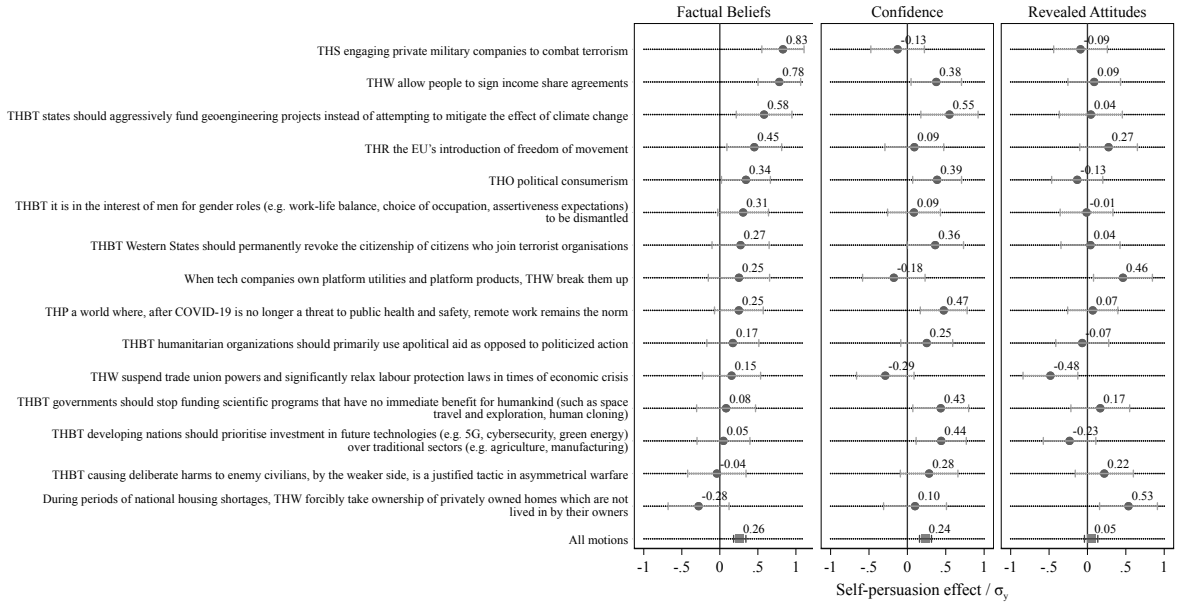


Figure 2: Self-persuasion by motion, ordered by effect size in Factual Beliefs

Notes: The figure shows self-persuasion effects β from regression model Equation (II.1), estimated separately by motion. The estimated effect is divided by the standard deviation in the outcome variable σ_y. Capped ranges are 95% confidence intervals. Acronyms: THS = This House supports, THW = This House would, THR = This House regrets, THBT = This House believes that, THO = This House opposes, THP = This House proposes.

These results are of interest for several reasons. First, it is striking that the debaters in our prestigious tournament succumb to the self-persuasion effect, while their (experienced) peers predict it. This suggests that self-persuasion works despite an awareness of its existence, as suggested in Saccardo and Serra-Garcia (2020). It also shows that people anticipate the biases of others, complementing results for the case of present-bias (Fedyk, 2018), and more conflicting evidence on overconfidence (Ludwig and Nafziger, 2011). Note that we asked adjudicators explicitly about their beliefs for both the opposition and opposition debaters, which may have raised the salience of this split. While it is hard to avoid such measurement effects, the anticipation of non-salient biases remains an open question. Second, the adjudictators seem able to use some of the content of the motions and questions in their predictions of the self-persuasion effect. This raises questions about the exact contextual features that generate self-persuasion. Given the limited number of motions/questions in our sample and the large number of potentially relevant dimensions, we leave this as a challenge for future research.

IV The why and how of self-persuasion

We now turn to the psychological mechanisms underlying self-persuasion and the potential benefits of self-persuasion. We first investigate a plausible mediator of the self-persuasion effect, namely the biased generation of arguments. We then discuss whether self-persuasion has instrumental value in helping debaters win the debate. We conclude this section by ruling out experimenter demand effects as a potential confound.

A The biased generation of arguments

A number of theories point to the biased generation of the number of arguments as a a mediator of self-persuasion. For instance, according to "persuasive argument theory" (Vinokur and Burstein, 1974), the number of new arguments that a side brings to the table is a key driver of persuasion. Mercier and Sperber (2011) theorize that our reasoning abilities have developed in order to persuade others through the biased generation of arguments, which produce self-persuasion as a by-product. Bénabou et al. (2019) argue that persuasion and justification in moral dilemmas occur as the result of a selective search for "narratives". On the empirical side, Thompson and Loewenstein (1992) shows that people asymmetrically recall facts in a bargaining situation, while other papers find people

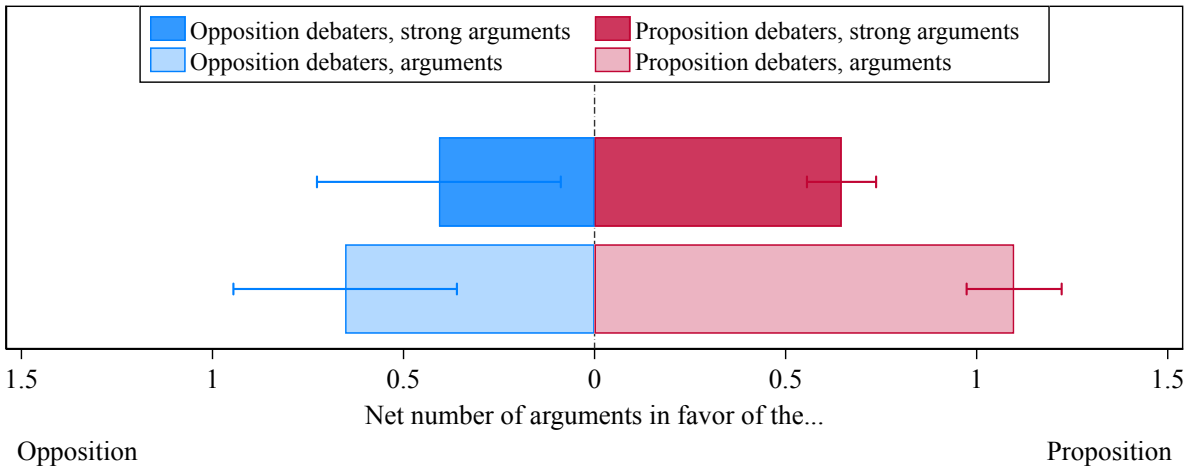


Figure 3: Differences in the number of arguments

Notes: The figure shows how debaters on each side disproportionately enumerate more arguments for their side at pre-debate. Bars with low intensity fill are based on all arguments debaters enumerate, and full bars are based on the number of arguments for each side that debaters deem as very strong. Capped ranges are 95% confidence intervals.

engage in biased search of evidence to align with their persuasion goals (Smith et al., 2017; Solda et al., 2019).

To understand the role of biased argument generation in our debate setting, we asked debaters in the pre-debate survey for the number of arguments they came up with during their preparation time, both for and against the motion. We also asked them how many of these arguments they considered to be "very strong". Figure 3 shows the average net number of arguments debaters came up with on both sides, split by treatment. As is clear from the graph, debaters engage in asymmetric selection of arguments. On average, they come up with almost one additional argument and one half of a "strong" argument in favor of their own side.

To quantify the impact of this asymmetry for self-persuasion, we conduct a parametric causal mediation analysis (Imai et al., 2010). We define s_i, the number of aligned arguments as a fraction of total arguments considered during preparation time, and investigate how this mediates self-persuasion on our three main outcome variable. The analysis decomposes the Average Treatment Effect (ATE) into the Average Direct Effect (ADE) and the Average Causal Effect Mediated by s_i (ACME). Overall, s_i drives 14 percent of the self-persuasion effect in Factual Beliefs, 44 percent for Confidence and 58 percent for Revealed Attitudes. In Online Appendix C we provide further details of the mediation effects both offline and online.

These results indicate a substantial, although heterogeneous and incomplete role for biased argu-

ment generation. To explain the effect of biased argument generation on beliefs, it must be the case that debaters fail to correct for this bias when they assess their position. This failure may either result from a lack of sophistication, in line with a literature on selection neglect (Juslin et al., 2007; Barron et al., 2019), or may itself be motivated by the wish to align beliefs and attitudes with the persuasion goal.

B Alignment and persuasiveness

Is self-persuasion "useful" to win a debate? The answer to this question speaks to theories about the social origins of self-persuasion. For instance, Von Hippel and Trivers (2011a) theorize that self-persuasion is a strategic action aimed at increasing persuasiveness through the reduction of nervous tics, give-away tells or other manifestations of cognitive dissonance that arise from a gap between beliefs and persuasion goals. This theory has received support in recent laboratory studies (Smith et al., 2017; Schwardmann and van der Weele, 2019; Solda et al., 2019).

While our experiment cannot distinguish this theory from a self-persuasion-as-byproduct account discussed above, we can test the prediction that self-persuasion has benefits for persuasion. To this end, we investigate the effect of having aligned beliefs on the "ballot score", a rating between 60 and 100 given by adjudicators to each individual debater at the end of the debate. The team with the highest ratings is declared the winner of the debate. To understand the usefulness of self-persuasion, we regress these scores on the alignment of Factual Beliefs and Confidence with the own persuasion goal, which are the variables which show substantial self-persuasion. Note that Factual Beliefs and Confidence at pre-debate are potentially endogenous as they depend partially on the individual degree of self-persuasion. Therefore, we also look at the alignment of Factual Beliefs in the baseline survey, which is exogenous due to the random assignment of persuasion goals.

Table 8 shows the results of these exercises. We find a positive correlation between alignment of beliefs and the ratings. For Factual Beliefs, the correlation is similar for both baseline beliefs (column 1) and pre-debate beliefs (column 2). For Confidence, where we don't have a baseline elicitation, we find a positive correlation with pre-debate beliefs (column 3). In all cases, the coefficients are marginally significant with $p < 0.100$, so this evidence is only indicative, and in need of confirmation by future research. In summary, we find some evidence in line with the idea that self-persuasion has a beneficial effect on the adjudicators' evaluations, which may explain its persistence in our sample of experienced debaters.

Table 8: Does alignment help to win the debate?

	Individual speaker score		
	(1)	(2)	(3)
Belief alignment with own side at baseline (standardized)	0.028	0.028	0.026
	(0.017)	(0.017)	(0.017)
Belief alignment with own side pre-debate (standardized)		0.032	
		(0.019)	
Confidence in own side pre-debate (standardized)			0.028
			(0.016)
Debaters	459	459	459
Observations	2179	2151	2148
R^2	0.029	0.033	0.031

Notes: Fixed effects linear regression model with standard errors (in parentheses) clustered at the team level. All specifications include motion and debater fixed effects. The number of observations is determined by valid responses from debaters over 5 (4 in Rotterdam) rounds of debate. The outcome is a metric of individual performance adjudicated in the ballot. Speaker scores for a handful of debaters could not be matched to our dataset, as they did not agree to this information becoming publicly available.

C Ruling out experimenter demand

Tappin et al. (2019) point out a common flaw in experiments that randomly assign persuasion goals to study politically motivated reasoning. Subjects may believe that the experimenter asked them to argue a particular viewpoint because of its empirical or logical validity, which aligns their beliefs even without any self-persuasion. The experimenter can avoid this by explicitly announcing the random nature of the assignment. However, this may lead the subject to second-guess the goal of the study, possibly introducing an experimenter demand effect.

Debating tournaments avoid these pitfalls, due to the nature of the randomization. Because it is public and explicit, debaters know not to infer anything from the assignment about the merits of their case. At the same time, it is a familiar and inconspicuous part of the competition and is therefore unlikely to direct participants' attention to our research question. To confirm this last claim, we asked subjects in the last survey to guess the aim of our research. We find that 19 percent of subjects made a guess that resembled our main hypotheses. If these subjects are driving our results, the effect should get smaller when we exclude them from the analysis. Online Appendix Table B.2 shows that this is not the case, indicating that experimenter demand is not a main factor in this setting.

V Conclusions

Our results show that the self-persuasion effects previously found in the laboratory are relevant in the field. We find that debaters distort factual beliefs and confidence in the direction of a position they

are randomly assigned to argue. Self-persuasion occurs despite incentives for accuracy and persists after an intense exposure to opposing views. These results obtain in prestigious tournaments, in a sample of experienced debaters that regularly supplies future elites and politicians.

Our result may contain insights for other applications, that we enumerate here. This extrapolation involves a degree of speculation, as there are alternative explanations that could be disentangled by future research. First, self-persuasion is likely to drive belief formation in political contexts, where convincing others is of central importance. This may explain why greater engagement with the political process causes greater and persistent polarization (Mullainathan and Washington, 2009) and why polarization is more severe in the US congress than it is in the American public (Fiorina and Abrams, 2008). It also suggests additional motives for political behavior such as canvassing and proselytizing, which may be important not just to convert others, but also for deepening the convictions of those doing the canvassing (Gal and Rucker, 2010).

Relatedly, self-persuasion may be at work in markets with asymmetric information. It predicts that sellers in economic transactions "drink the kool-aid" and become overly optimistic about their product. This may explain why financial advisors privately invest in the under-performing funds for which they receive sales commissions (Linnainmaa et al., 2018). It may also be a driving force behind the development of asset market bubbles, for instance during the financial crisis of 2007-8, where private real-estate portfolios of agents working in sales departments of mortgage providers under-performed those of other agents as well as non-specialists (Cheng et al., 2015). Self-persuasion may also be involved in the spectacular rise and fall of start-up companies like Theranos, as entrepreneurs trying to lure investors become overconfident and miscalibrated.

Finally, one may wonder if there is a connection between our findings and polarization among "regular" people. Unlike politicians, lawyers and entrepreneurs, most people do not earn money for persuading others. Yet, as evidenced by heated discussions on social media, at dinner tables and at football games, many people are intrinsically motivated to convince others of what they believe to be true or what aligns with their identity. In our setting, we cannot disentangle the relative importance of intrinsic and extrinsic motives to be persuasive, as tournament debaters are likely driven by both. On the one hand, they are engaged in a quest for status and visibility. On the other hand, they are unpaid enthusiasts who enjoy the act of persuasion. We conjecture that both types of motivation can induce self-persuasion including among non-professionals, but testing this conjecture remains a task for future research.

Our results leave open some other interesting questions. For instance, measuring the impact of persistent or long-run persuasion goals, like those arising from group membership or party affiliation, may help understand the formation of personal identity. Another question concerns the design of institutions that revolve around debating. We show that debates do not necessarily resolve conflicts of opinion and can actually make them worse. At the same time, debating tournaments are an extremely competitive context, and our results may not extend to settings where parties aim to reach consensus (Felton et al., 2015). Thus, an important question is how to design debating contexts to promote a shared understanding of facts and mitigate disagreement.

References

Andrews, Isaiah and Maximilian Kasy, "Identification of and correction for publication bias," *American Economic Review*, 2019, *109* (8), 2766–94.

Babcock, Linda and George Loewenstein, "Explaining Bargaining Impasse: The Role of Self-Serving Biases," *Journal of Economic Perspectives*, March 1997, *11* (1), 109–126.

_ , _ , Samuel Issacharoff, and Colin Camerer, "Biased Judgments of Fairness in Bargaining," *The American Economic Review*, 1995, *85* (5), 1337–1343.

Bail, Christopher A., Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, M. B. Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky, "Exposure to opposing views on social media can increase political polarization," *Proceedings of the National Academy of Sciences*, 2018, *115* (37), 9216–9221.

Barron, Kai, Steffen Huck, and Philippe Jehiel, "Everyday econometricians: Selection neglect and overoptimism when learning from others," Technical Report, WZB Discussion Paper 2019.

Bénabou, Roland and Jean Tirole, "Mindful Economics: The Production, Consumption, and Value of Beliefs," *Journal of Economic Perspectives*, September 2016, *30* (3), 141–164.

Bénabou, Roland, Armin Falk, and Jean Tirole, "Narratives, Imperatives and Moral Reasoning," 2019.

Bullock, John G, Alan S Gerber, Seth J Hill, Gregory A Huber et al., "Partisan Bias in Factual Beliefs about Politics," *Quarterly Journal of Political Science*, 2015, *10* (4), 519–578.

Chen, Zhuoqiong Charlie and Tobias Gesche, "Persistent bias in advice-giving," *University of Zurich, Department of Economics, Working Paper*, 2017, (228).

Cheng, Ing-Haw, Sahil Raina, and Wei Xiong, "Wall Street and the Housing Bubble: Bad Incentives, Bad Models, or Bad Luck?," *American Economic Review*, 2015, *104* (9), 2797–2829.

Cochran, William G, "The combination of estimates from different experiments," *Biometrics*, 1954, *10* (1), 101–129.

Coutts, Alexander, "Testing Models of Belief Bias: An Experiment," *Games and Economic Behavior*, 2019, *113*, 549–565.

DellaVigna, Stefano and Devin Pope, "Predicting experimental results: who knows what?," *Journal of Political Economy*, 2018, *126* (6), 2410–2456.

_ **and Elizabeth Linos**, "Rcts to scale: Comprehensive evidence from two nudge units," Technical Report, National Bureau of Economic Research 2020.

_ **, Nicholas Otis, and Eva Vivalt**, "Forecasting the Results of Experiments: Piloting an Elicitation Strategy," in "AEA Papers and Proceedings," Vol. 110 2020, pp. 75–79.

Di Tella, Rafael, S. Galiant, and E. Schargrodsky, "The Formation of Beliefs: Evidence from the Allocation of Land Titles to Squatters," *The Quarterly Journal of Economics*, 2007, *122* (1), 209–241.

Elliot, Andrew J and Patricia G Devine, "On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort.," *Journal of personality and social psychology*, 1994, *67* (3), 382.

Engel, Christoph and Andreas Glöckner, "Role-Induced Bias in Court: An Experimental Analysis," *Journal of Behavioral Decision Making*, 2013, *26* (3), 272–284.

Fedyk, Anastassia, "Asymmetric naivete: Beliefs about self-control," *Available at SSRN 2727499*, 2018.

Felton, Mark, Amanda Crowell, and Tina Liu, "Arguing to agree: Mitigating my-side bias through consensus-seeking dialogue," *Written Communication*, 2015, *32* (3), 317–331.

Festinger, Leon and James M Carlsmith, "Cognitive consequences of forced compliance.," *The journal of abnormal and social psychology*, 1959, *58* (2), 203.

_ **and James M. Carlsmith**, "Cognitive consequences of forced compliance," *Journal of abnormal psychology*, 1959, *58* (2), 203–210.

Fiorina, Morris P. and Samuel J. Abrams, "Political polarization in the American public," *Annual Revue of Political Science*, 2008, *11*, 563–588.

Gal, David and Derek D Rucker, "When in doubt, shout! Paradoxical influences of doubt on proselytizing," *Psychological Science*, 2010, *21* (11), 1701–1707.

Ganguly, Ananda and Joshua Tasoff, "Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future," *Management Science*, 2016, (September 2017), mnsc.2016.2550.

Gentzkow, Matthew, "Polarization in 2016," *Toulouse Network for Information Technology Whitepaper*, 2016, pp. 1–23.

Gneezy, Uri, Silvia Saccardo, Marta Serra-Garcia, and Roel van Veldhuizen, "Bribing the self," *Games and Economic Behavior*, 2020, *120*, 311–324.

Higgins, Julian PT and Simon G Thompson, "Quantifying heterogeneity in a meta-analysis," *Statistics in medicine*, 2002, *21* (11), 1539–1558.

Hippel, Svenja and Sven Hoeppner, "Biased judgements of fairness in bargaining: A replication in the laboratory," *International Review of Law and Economics*, 2019, *58*, 63–74.

Hippel, William Von and Robert Trivers, "The evolution and psychology of self-deception," *Behavioral and Brain Sciences*, 2011, *34* (1), 1–16.

Hossain, Tanjim and Ryo Okui, "The binarized scoring rule," *Review of Economic Studies*, 2013, *80* (3), 984–1001.

Huffman, David, Collin Raymond, and Julia Shvets, "Persistent overconfidence and biased memory: Evidence from managers," *Pittsburgh: University of Pittsburgh*, 2019.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto, "Identification, inference and sensitivity analysis for causal mediation effects," *Statistical science*, 2010, pp. 51–71.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J Westwood, "The origins and consequences of affective polarization in the United States," *Annual Review of Political Science*, 2019, *22*, 129–146.

Janis, Irving L. and Bert T. King, "The influence of role playing on opinion change," *Journal of Abnormal and Social Psychology*, 1954, *49* (2), 211–218.

Juslin, Peter, Anders Winman, and Patrik Hansson, "The naive intuitive statistician: a naive sampling model of intuitive confidence intervals.," *Psychological review*, 2007, *114* (3), 678.

Levitt, Steven D and John A List, "What do laboratory experiments measuring social preferences reveal about the real world?," *Journal of Economic perspectives*, 2007, *21* (2), 153–174.

Levy, Ro'ee, "Social media, news consumption, and polarization: Evidence from a field experiment," *American economic review*, 2021, *111* (3), 831–70.

Linnainmaa, Juhani T, Brian Melzer, and Alessandro Previtero, "The misguided beliefs of financial advisors," *Kelley School of Business Research Paper*, 2018, (18-9).

List, John A, "Does market experience eliminate market anomalies?," *The Quarterly Journal of Economics*, 2003, *118* (1), 41–71.

_ , "The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions," *Journal of political Economy*, 2006, *114* (1), 1–37.

Loewenstein, George, Samuel Issacharoff, Colin Camerer, and Linda Babcock, "Self-Serving Assessments of Fairness and Pretrial Bargaining," *The Journal of Legal Studies*, January 1993, *22* (1), 135–159.

Lord, Charles G, Lee Ross, and Mark R. Lepper, "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence.," *Journal of Personality and Social Psychology*, 1979, *37* (11), 2098.

_ **, Mark R Lepper, and Elizabeth Preston**, "Considering the opposite: a corrective strategy for social judgment.," *Journal of personality and social psychology*, 1984, *47* (6), 1231.

Ludwig, Sandra and Julia Nafziger, "Beliefs about overconfidence," *Theory and decision*, 2011, *70* (4), 475–500.

Mayraz, Guy, "Wishful Thinking," *CEP Discussion Paper*, 2011, (1092).

Mercier, Hugo, "Self-deception: adaptation or by-product? (Comment on Von Hippel and Trivers)," *Behavioral and Brain Sciences*, 2011, *34* (1), 35.

_ **and Dan Sperber**, "Why do humans reason? Arguments for an argumentative theory," *Behavioral and brain sciences*, 2011, *34* (2), 57–74.

Mullainathan, Sendhil and Ebonya Washington, "Sticking with your vote: Cognitive dissonance and political attitudes," *American Economic Journal: Applied Economics*, 2009, *1* (1), 86–111.

O'Neill, Patrick and Diane E. Levings, "Inducing biased scanning in a group setting to change attitudes toward bilingualism and capital punishment.," *Journal of Personality and Social Psychology*, 1979, *37* (8), 1432–1438.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey, "Optimal expectations and limited medical testing: Evidence from huntington disease," *American Economic Review*, 2013, *103* (2), 804–830.

Perkins, David, "Learning to reason: The influence of instruction, prompts and scaffolding, metacognitive knowledge, and general intelligence on informal reasoning about everyday social and political issues," *Judgment and Decision Making*, 2019, *14* (6), 624.

Saccardo, Silvia and Marta Serra-Garcia, "Cognitive Flexibility or Moral Commitment? Evidence of Anticipated Belief Distortion," 2020.

Schlag, Karl H. and Joël J. Van der Weele, "Eliciting Probabilities, Means, Medians, Variances and Covariances without Assuming Risk Neutrality," *Theoretical Economics Letters*, 2013, *3* (1), 38–42.

Schwardmann, Peter and Joël J. van der Weele, "Deception and Self-Deception," *Nature Human Behavior*, 2019, *3*, 1055–1061.

_ , Egon Tripodi, and Joël van der Weele, "Self-Persuasion: Evidence from Field Experiments at Two International Debating Competitions," 2019.

Smith, Megan K., Robert Trivers, and William von Hippel, "Self-deception facilitates interpersonal persuasion," *Journal of Economic Psychology*, December 2017, *63*, 93–101.

Solda, Alice, Changxia Ke, Lionel Page, and William Von Hippel, "Strategically delusional," *Experimental Economics*, 2019, pp. 1–28.

Taber, Charles S. and Milton Lodge, "Motivated Skepticism in the Evaluation of Political Beliefs," *American Journal of Political Science*, 2006, *50* (3), 755–769.

Tappin, Ben M, "Estimating the between-issue variation in party elite cue effects," 2020.

_ , Gordon Pennycook, and David Rand, "Thinking clearly about causal inferences of politically motivated reasoning: Why paradigmatic study designs often prevent causal inference," 2019.

Thompson, Leigh and George Loewenstein, "Egocentric interpretations of fairness and interpersonal conflict," *Organizational Behavior and Human Decision Processes*, 1992, *51* (2), 176–197.

Vinokur, Amiram and Eugene Burstein, "Effects of partially shared persuasive arguments on group-induced shifts: A group-problem-solving approach," *Journal of Personality and Social Psychology*, 1974, *29* (3), 305–315.

Von Hippel, William and Robert Trivers, "The evolution and psychology of self-deception," *The Behavioral and brain sciences*, 2011, *34* (1), 1–16; discussion 16–56.

Zimmermann, Florian, "The dynamics of motivated beliefs," *American Economic Review*, 2020, *110* (2), 337–61.

Online Appendix for "Self-Persuasion: Evidence from Field Experiments at International Debating Competitions"*

Peter Schwardmann Egon Tripodi Joël J. van der Weele

November 4, 2021

A Additional tables and figures

Table A.1: Debaters' responsibilities by role

Team	Speaking role	Speaking order	Team	Speaking role	Speaking order
Opening Government (OG)	Prime Minister (PM)	First	Opening Opposition (OO)	Leader of the Opposition (LO)	Second
	• Defines and interprets the motion • Develops the case for the proposition			• Accepts definition of the motion • Refutes the case of OG • Constructs arguments against PM's interpretation of the motion	
	Deputy Prime Minister (DPM)	Third		Deputy Leader of the Opposition (DLO)	Fourth
	• Refutes the case of OO • Rebuilds the case of OG • May add new arguments to the case of the PM			• Continues refuting the case of OG • Rebuilds the case of OO • May add new arguments to the case of the LO	
Closing Government (CG)	Member of the Government (MG)	Fifth	Closing Opposition (CO)	Member of the Opposition (MO)	Sixth
	• Defends the general direction and case of OG • Continues refutation of OO • Develops a new argument that is different from but consistent with the case of OG			• Defends the general direction taken by OO • Continues general refutation of OG's case • Provides more specific refutation of CG's case • Provides new opposition arguments	
	Government Whip (GW)	Seventh		Opposition Whip (OW)	Eighth
	• Summarizes the entire debate from the point of view of the proposition, defending the general view point of both OG and CG with a special eye toward the case of CG • Does not provide new arguments			• Summarizes the entire debate from the point of view of the opposition, defending the general view point of both OO and CO with a special eye toward the case of CO • Does not provide new arguments	

*Schwardmann: Carnegie Mellon University; schwardmann@cmu.edu. Tripodi: University of Essex and JILAEE; egon.tripodi@essex.ac.uk. van der Weele: University of Amsterdam and Tinbergen Institute; j.j.vanderweele@uva.nl.

Table A.2: Debaters' baseline belief alignment and characteristics, by side of the motion

Sample:	Full sample	Opposition	Proposition	p-value
Baseline belief alignment	50.177	50.111	50.243	0.928
	(0.638)	(0.906)	(0.898)	
Female	0.348	0.352	0.345	0.731
	(0.010)	(0.014)	(0.014)	
Age	21.499	21.474	21.525	0.630
	(0.067)	(0.094)	(0.096)	
Time in debating	2.467	2.453	2.481	0.503
	(0.022)	(0.031)	(0.031)	
Past achievements	7.524	7.454	7.594	0.650
	(0.275)	(0.389)	(0.389)	
Local nationality	0.148	0.153	0.144	0.546
	(0.008)	(0.011)	(0.010)	
Left to right (0-10) political ideology scale	3.515	3.537	3.493	0.487
	(0.041)	(0.058)	(0.058)	
Unique debaters	462	233	229	
Observations	2238	1119	1119	

Notes: P-value is from a one-way ANOVA on ranks (Kruskal-Wallis) test comparing the two groups.

Table A.3: Self-persuasion, robustness to clustering at the room level

Time of elicitation:	Pre-debate			Post-debate		
	(1)	(2)	(3)	(4)	(5)	(6)
Proposition alignment in:	Factual Beliefs	Confidence	Revealed Attitudes	Factual Beliefs	Confidence	Revealed Attitudes
Assigned to proposition	7.157	5.795	0.097	5.055	7.963	0.200
	(1.184)	(1.075)	(0.100)	(1.289)	(1.378)	(0.093)
Debaters	473	473	473	470	274	470
Observations	2217	2213	2212	2171	1286	2169
R^2	0.213	0.104	0.199	0.227	0.111	0.228

Notes: Pooled ordinary least square model. Standard errors in parentheses are clustered at the room level, which varies at the team level over time. All specifications include question fixed effects. Each round, debaters are randomly assigned to argue either as proposition or opposition. The outcome of columns 1-3 is our measure of post-debate alignment with the proposition in either Factual Beliefs, Confidence, or Revealed Attitudes. The outcome of columns 4-6 is our measure of alignment with the proposition in either Factual Beliefs, Confidence, or Revealed Attitudes–either at pre-debate or post-debate. For all outcomes, higher values denote greater alignment with the proposition. The number of observations is determined by valid responses from debaters over 5 (4 in Rotterdam) rounds of debate. Post-debate Confidence was collected only at online tournaments.

Table A.4: Pre-debate self-persuasion, robustness to the omission of fixed effects

Sample:	All tournaments			Munich and Rotterdam (Offline)			Amsterdam and LSE (Online)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Proposition alignment in:	Factual Beliefs	Confidence	Revealed Attitudes	Factual Beliefs	Confidence	Revealed Attitudes	Factual Beliefs	Confidence	Revealed Attitudes
Assigned to proposition	7.727	5.982	0.114	5.900	4.381	0.300	8.938	7.036	-0.005
	(1.301)	(1.014)	(0.121)	(2.115)	(1.504)	(0.150)	(1.644)	(1.359)	(0.174)
Debaters	473	473	473	196	196	196	277	277	277
Observations	2217	2213	2212	884	883	883	1333	1330	1329
R^2	0.021	0.024	0.000	0.012	0.014	0.006	0.027	0.032	0.000

Notes: Random effects linear regression model with standard errors (in parentheses) clustered at the team level. This table replicates the analysis in Table 2 excluding question fixed effects. Each round, debaters are randomly assigned to argue either as proposition or opposition. The outcome is our measure of pre-debate alignment with the proposition in either Factual Beliefs, Confidence, or Revealed Attitudes. For all three outcomes, higher values denote greater alignment with the proposition. The support of Factual beliefs and Confidence includes integers between 0 and 100, while Revealed Attitudes includes integers between -4 and 4. The number of observations is determined by valid responses from debaters over 5 (4 in Rotterdam) rounds of debate.

Table A.5: Political polarization

Time of elicitation:	Pre-debate			Post-debate		
	(1)	(2)	(3)	(4)	(5)	(6)
Proposition alignment in:	Factual Beliefs	Confidence	Revealed Attitudes	Factual Beliefs	Confidence	Revealed Attitudes
	Panel a). Excluding neutral motions					
Politically aligned with proposition	3.858	0.426	0.468	0.348	-2.492	0.271
	(1.504)	(1.358)	(0.129)	(1.472)	(1.663)	(0.135)
Debaters	462	462	462	461	270	461
Observations	1786	1782	1781	1756	892	1754
R^2	0.182	0.094	0.240	0.238	0.163	0.236
	Panel b). All motions					
Politically aligned with proposition	2.386	0.945	0.310	0.178	-1.600	0.200
	(1.269)	(1.278)	(0.111)	(1.283)	(1.525)	(0.117)
Debaters	463	463	463	462	271	462
Observations	2178	2174	2173	2141	1277	2139
R^2	0.199	0.086	0.206	0.233	0.123	0.227

Notes: Random effects linear regression model with standard errors (in parentheses) clustered at the team level. All specifications include question fixed effects. The outcome in columns 1-3 is our measure of pre-debate alignment with the proposition in either Factual Beliefs, confidence, or Revealed Attitudes. The outcome in columns 4-6 is the post-debate alignment analog. The number of observations is determined by valid responses from debaters over 5 (4 in Rotterdam) rounds of debate. This categorization is based on a small survey we did with 23 debaters in April 2021 to assess the robustness of results in Table 4. Motions were categorized as being left or right wing on a 5 points scale. We call a motion left (right) leaning if the average rating from debaters is below (above) 3. In panel a we exclude neutral motions from the analysis (average rating between 2.9 and 3.1).

Table A.6: Self-persuasion on day two

	Proposition alignment in Factual Beliefs	
	(1)	(2)
Sample:	All motions	Excluding round 5
Assigned to proposition	4.055	4.027
	(1.361)	(1.519)
Debaters	254	254
Observations	1266	1012
R^2	0.297	0.182

Notes: Random effects linear regression model with standard errors (in parentheses) clustered at the team level. All specifications include question fixed effects. The outcome is our measure of alignment with the proposition in Factual Beliefs obtained from asking debaters on day 2 of the competition the same 5 factual questions related to the motions that they were asked at baseline. Higher values denote greater alignment with the proposition. This data was collected only at the online competitions. Column 2 excludes the fifth round of each tournament, which is the only one that takes place on day 2. The number of observations is determined by valid responses from debaters over 5 (4 in column 2) rounds of debate.

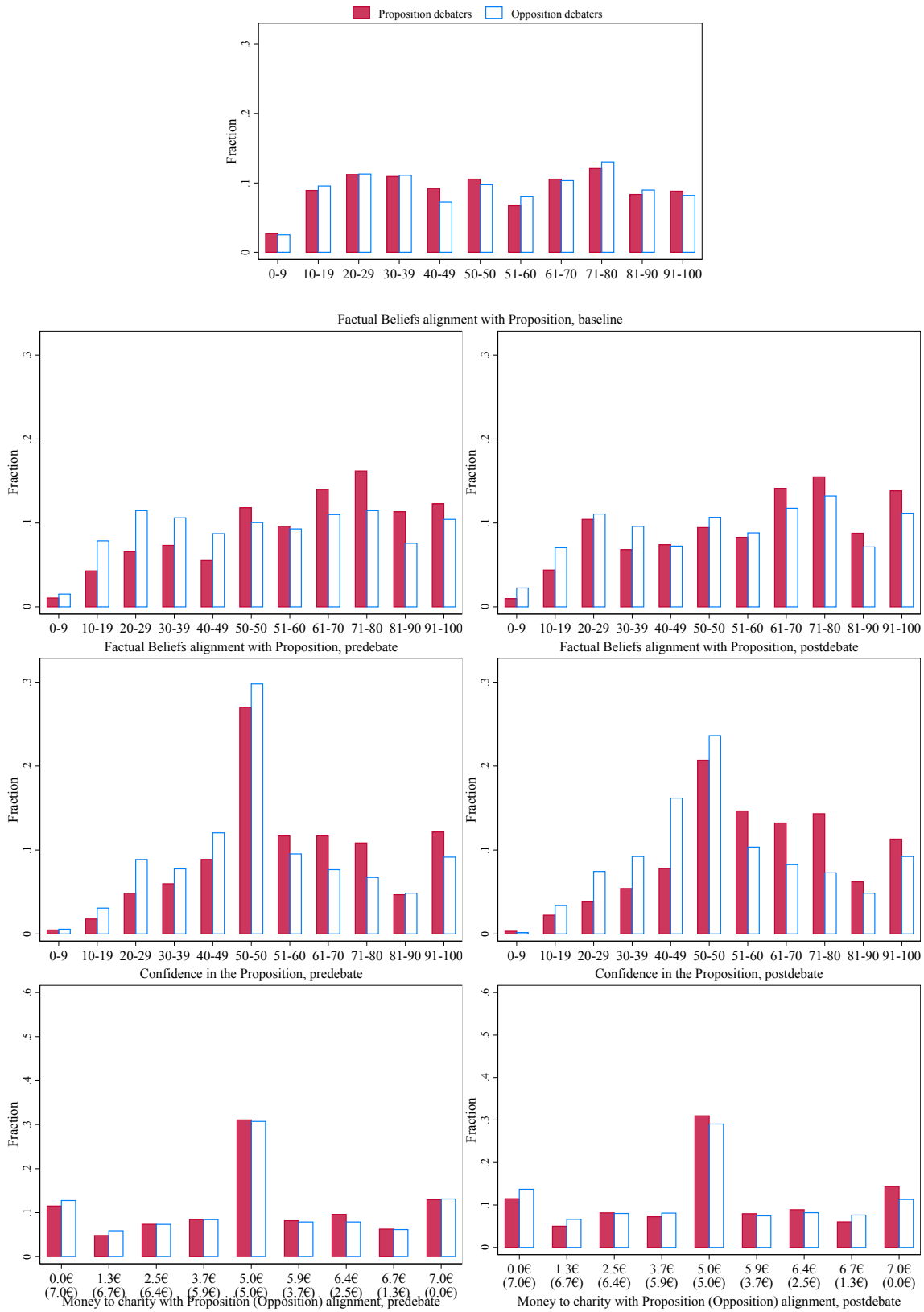


Figure A.1: Distribution of alignment

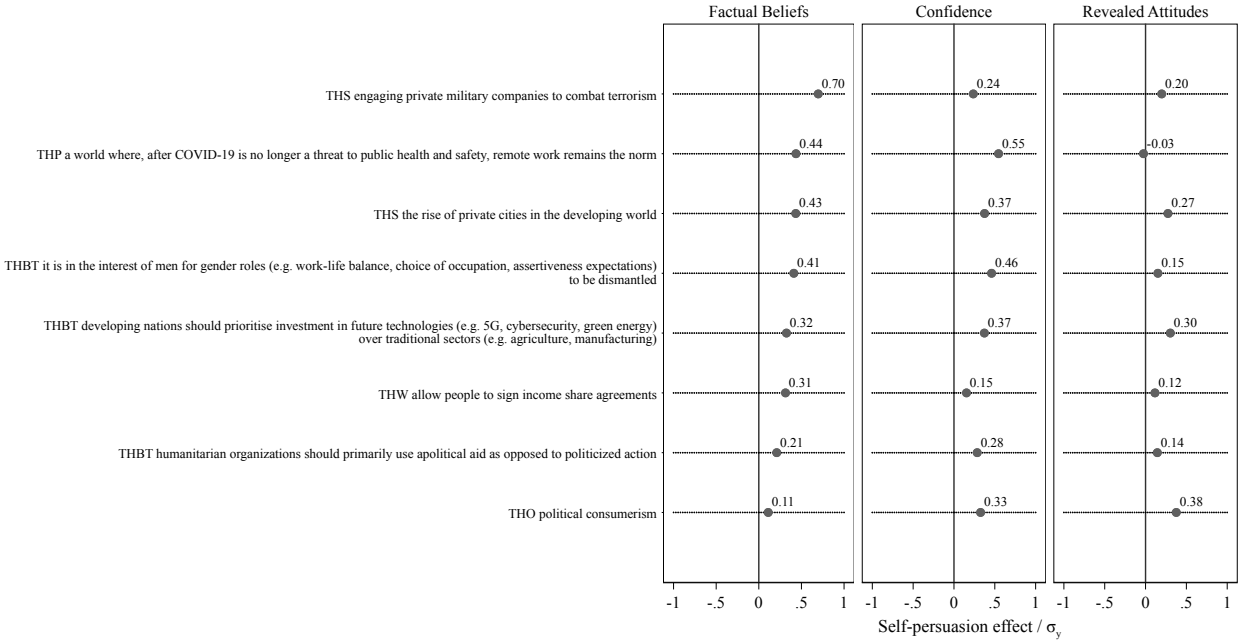


Figure A.2: Self-persuasion adjudicators' predictions by motion, ordered by predicted effect size in Factual Beliefs. We show only 8 out of 10 online motions, due to some of the motions' sensitive nature. Our analysis is based on all 10 online motions.

B Robustness to experimenter demand

Subjects who infer the research hypotheses under investigation may distort their reports to help the researchers prove their hypotheses, or answer in a way that they think is expected of them. In our study, incentives for accuracy should mitigate such distortions, but may not entirely eliminate them. We therefore asked subjects in the final survey of our experiment to write down in an open field text box what they thought the research was trying to demonstrate.

Table B.1: Categorization of debaters' response

Persuasion incentives cause self-persuasion	0.187
Alignment helps being persuasive	0.040
Better informed debaters are more persuasive	0.075
Views change through the debate	0.244
Overly generic answer	0.292
Other research questions	0.312
Answered question	401
Observations	473

Notes: Open-field answers are categorized by a research assistant; each answer may be assigned multiple categories when the respondent guesses multiple research questions. We report shares of respondents in each category among the 85 percent of debaters in our sample who answered this open-field question.

In Table B.1 we report the result of a manual categorization of non-blank responses (85 percent of the sample). Among these, 29 percent give a generic answer (e.g. "Game-theory"), while the rest seem to have in mind some concrete research hypotheses. The majority of subjects reported fairly sophisticated guesses.[1] The most frequent remaining categories are our residual category "Other research questions", that includes questions that were not part of our pre-registered hypotheses (e.g. how knowledgeable debaters are compared to the general population, how salience of authority affects trust in factual evidence, whether greater availability of factual arguments increases persuasiveness). Almost 20 percent of subjects guessed the primary research hypothesis of self-persuasion.

If experimenter demand drives our treatment effect, our effects should go down when we exclude those who correctly guessed the effect of the research. In Table B.2 we therefore re-estimate our main result without this category of subjects. Reassuringly, we find that the magnitudes of the differences in

[1] Some responses were fairly accurate in capturing many of the research hypotheses (e.g. *"1. See how engaging with motion from a certain assigned point of view influences perception of facts in accordance to position in debate 2. how belief/being convinced of position in debate affects debaters persuasiveness (that's why you gave us scores on persuasion and rhetoric as well) –> How debating from assigned point of view affects opinion and how that affects performance in debate"*, some others completely miss the main hypotheses (e.g. *"Connection between knowledge and persuasiveness? - Not sure, would love to find out!"*), and some others are overly generic (e.g. *"Game-theory"*, *"Curse the competition"*).

all three outcomes between proposition and opposition speakers, estimated for the subset of "unaware subjects", are very similar to the ones estimated in the full sample, indicating that demand effects do not drive our results.[2]

Table B.2: Replication of main results excluding subjects who could guess the research hypothesis at the end of the tournament

	Pre-debate			Post-debate		
Proposition alignment in:	(1) Factual Belief	(2) Confidence	(3) Revealed Attitudes	(4) Factual Belief	(5) Confidence	(6) Revealed Attitudes
Assigned to proposition	6.948	5.732	0.0491	3.901	8.378	0.150
	(1.144)	(1.070)	(0.102)	(1.443)	(1.445)	(0.103)
Debaters	398	398	398	395	229	395
Observations	1856	1852	1851	1813	1068	1811
R^2	0.226	0.104	0.183	0.227	0.158	0.206

Notes: Column 1 and 4 replicate analysis in column 1 of Table 2 and Table 5 excluding subjects who guessed the research hypothesis of self-persuasion on facts. Column 2 and 5 replicate analysis in column 2 of Table 2 and Table 5 excluding subjects who guessed the research hypothesis that debaters who be relatively more confident of the merits of their own position. Column 3 and 6 replicates analysis in column 3 of Table 2 and Table 5 excluding subjects who guessed the research hypothesis of self-persuasion on the values of social causes.

C Mediation

Our discussion proposes that persuasion goals can have both a direct effect on belief alignment due to strategic choice of beliefs and an indirect effect due to the cognitive constraints that generate bias when debaters sample an unbalanced set of arguments to prepare their speech. In a linear framework, such direct and indirect effects can be assessed through the following system of structural equations

$$\begin{aligned} Y_i &= \alpha_1 + \beta_1 T_i + \phi_1 X_i + \epsilon_{i1} \\ M_i &= \alpha_2 + \beta_2 T_i + \phi_2 X_i + \epsilon_{i2} \\ Y_i &= \alpha_3 + \beta_3 T_i + \gamma M_i + \phi_3 X_i + \epsilon_{i3} \end{aligned} \tag{C.1}$$

where standard notation is used for expositional purposes: Y_i is the outcome of interest, T_i is the treatment variable, M_i is the intermediate outcome measure after treatment that mediates the treatment effect, and X_i is a vector of controls. β_1 represents the average treatment effect (ATE), which

[2] De Quidt et al. (2018) propose to estimate bounds on the impact of experimenter demand effects through the use of additional treatments that increase awareness among subjects of the experimenters' research hypotheses De Quidt et al. (2018). Providing such exogenous variation of awareness of research hypotheses was not feasible in the context of our field experiment.

includes both direct and indirect effects of the treatment on the main outcome of interest. If the structural equations are correctly specified, a *sequential ignorability* assumption allows us to interpret $\gamma\beta_2$ as the *causal* indirect effect of T_i, mediated through M_i, on Y_i Imai et al. (2010b).

Sequential ignorability requires that (i) conditional on X_i, the potential values of the outcome and the mediator are distributed independently of the treatment, and (ii) conditional on T_i and X_i, the potential outcome is distributed independently of the observed mediator. Both conditions are fairly strong. Because our treatment assignment is randomized, the first condition is met by design. However, the second condition does not directly follow from random assignment, and is hard to test. If the second condition is met, we would expect that the outcome and the mediator are uncorrelated within treatment. Figure C.1 provides supporting evidence of the lack of such correlation.

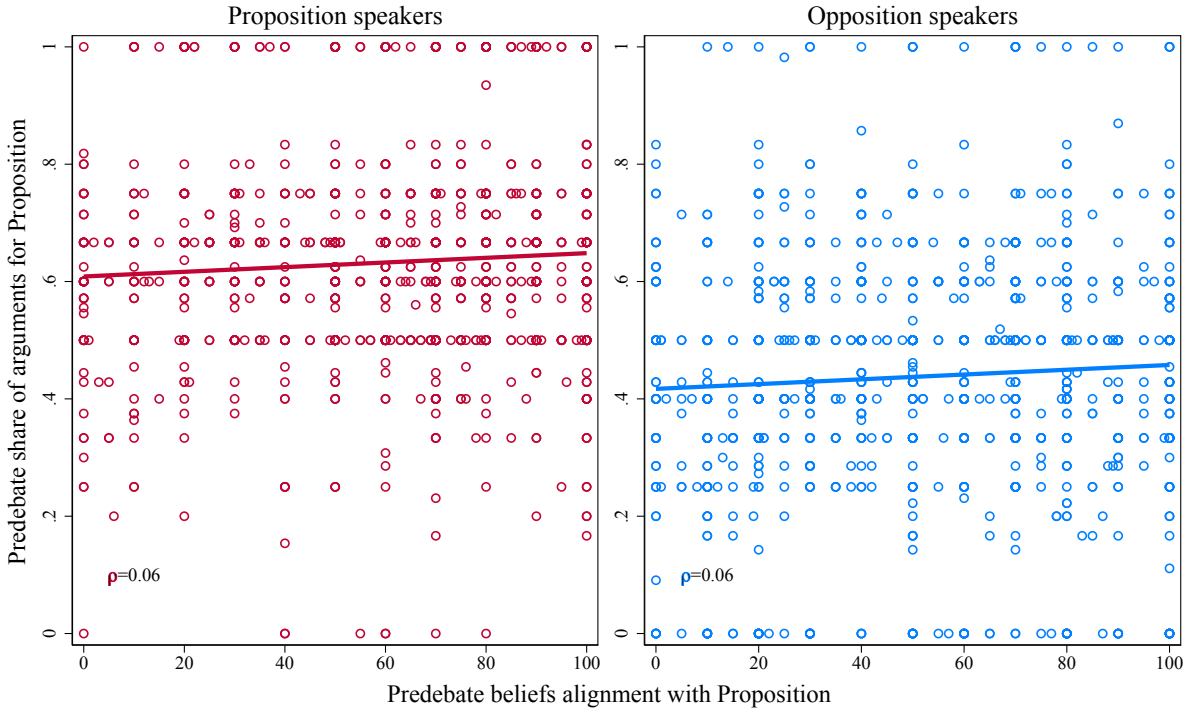


Figure C.1: Correlation between share of proposition arguments and pre-debate belief alignment, within each side of the debate

In Figure C.2 we include diagrams that illustrate potential causal links between the treatment, mediating factors, and the outcome. Assuming sequential ignorability rules out causal links between mediators (sub-figures (d) and (e)), but allows for multiple downstream causal relationships from treatment, through mediators, to the outcome of interest (sub-figures (a) to (c)), so that by estimating $\gamma\beta_2$ from C.1 we could directly obtain a valid estimate of the causal effect of the treatment mediated

through M_i.

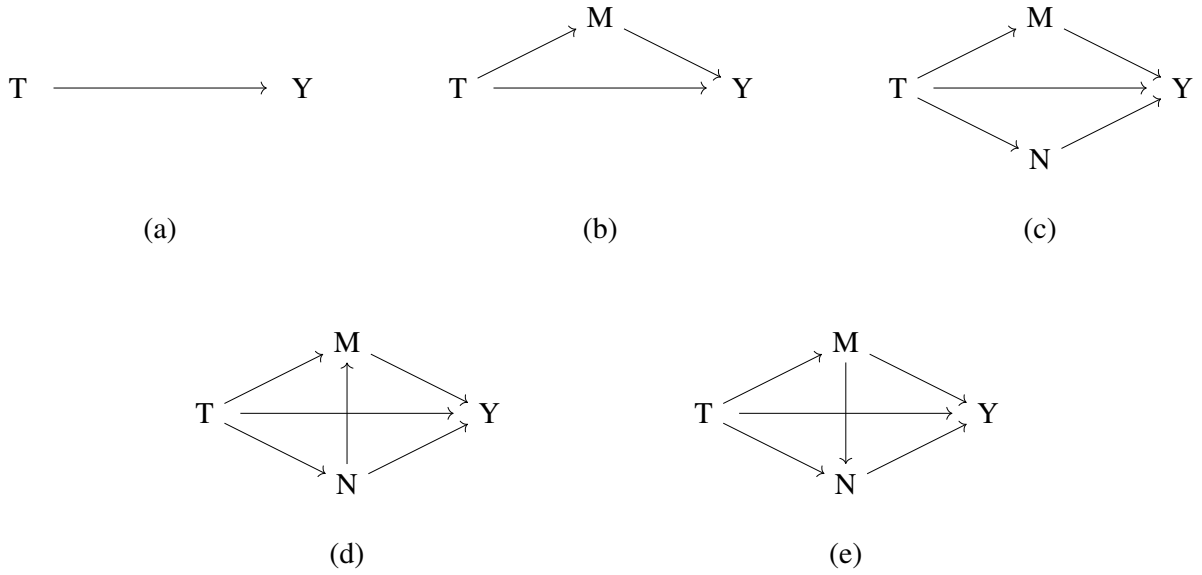


Figure C.2: Diagrams representing possible causal mechanisms between treatment, mediating outcomes, and main outcome

Notes: In (a), the outcome can only be affected directly by the treatment variable. In (b), the treatment affects both the outcome directly and an intermediate mediator; the mediator in turn affects the outcome. In (c), the treatment affects both the outcome directly and two intermediate mediators; both mediators in turn affect the outcome. In (d) and (e), the treatment affects both the outcome directly and two intermediate mediators; both mediators in turn affect the outcome, and mediators also affect one another.

In the potential outcome framework with binary treatment $t \in \{0,1\}$ and one mediator it is straightforward to derive the causal mediated effect directly as a component of the average treatment effect $\tau_i = Y_i(1) - Y_i(0)$, which can be equivalently written as $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$. With some algebra, it is simple to obtain that

$$2[Y_i(1, M_i(1)) - Y_i(0, M_i(0))] = \overbrace{Y_i(1, M_i(1)) - Y_i(1, M_i(0))}^{\delta_i(1)} + \overbrace{Y_i(0, M_i(1)) - Y_i(0, M_i(0))}^{\delta_i(0)} +$$
$$+ \overbrace{Y_i(1, M_i(1)) - Y_i(0, M_i(1))}^{\zeta_i(1)} + \overbrace{Y_i(1, M_i(0)) - Y_i(0, M_i(0))}^{\zeta_i(0)}$$

where $\delta(t)$ defines the indirect effect of the treatment in treatment t, and $\zeta_i(t)$ defines the direct effect of the treatment holding constant the level of the mediator at the treatment t level. When $\delta_i(t) = \delta_i$ and $\zeta_i(t) = \zeta_i$ for any t, there is no interaction between treatment and mediator, and the ATE can

simply be expressed as $\tau_i = \delta_i + \zeta_i$, yielding a simple decomposition of the ATE in average causal mediated effect (ACME) and average direct effect (ADE).

To identify the ACME of persuasion goals on belief alignment with proposition b_i through the share of proposition arguments considered during preparation period s_i, we estimate the following random effects models with standard errors clustered at the team level

$$
\begin{aligned}
&\textit{Model 1:} && b_{i,q} = \alpha_1 + \beta_1 proposition_{i,q} + \phi_1 X_i + \epsilon_{i1,q} \\
&\textit{Model 2:} && s_{i,q} = \alpha_2 + \beta_2 proposition_{i,q} + \phi_2 X_i + \epsilon_{i2,q} \\
&\textit{Model 3:} && b_{i,q} = \alpha_3 + \beta_3 proposition_{i,q} + \gamma s_{i,q} + \phi_3 X_i + \epsilon_{i3,q}
\end{aligned}
$$

and use sampling distributions of the parameter estimates from *model 1* to simulate potential outcomes $b_{i,q}(proposition_{i,q} = 1)$ and $b_{i,q}(proposition_{i,q} = 0)$, from *model 2* to simulate potential outcomes $s_{i,q}(proposition_{i,q} = 1)$ and $s_{i,q}(proposition_{i,q} = 0)$, and from *model 3* to simulate potential outcomes $b_{i,q}(1, s_{i,q}(1))$, $b_{i,q}(0, s_{i,q}(1))$, $b_{i,q}(1, s_{i,q}(0))$, and $b_{i,q}(0, s_{i,q}(0))$. Table C.1 reports the results from this exercise.

Table C.1: Decomposition of pre-debate treatment effect in mediated and direct effect

Sample:	All tournaments			Offline			Online		
Proposition alignment in:	Factual Beliefs	Confidence	Revealed Attitudes	Factual Beliefs	Confidence	Revealed Attitudes	Factual Beliefs	Confidence	Revealed Attitudes
Average causal effect mediated by s_i (ACME)	1.031	2.614	0.066	1.793	2.303	0.133	0.749	3.157	0.007
	(0.653)	(0.560)	(0.048)	(1.230)	(1.084)	(0.079)	(0.617)	(0.576)	(0.057)
Average direct effect (ADE)	6.179	3.277	0.046	4.612	2.046	0.165	7.008	3.749	-0.019
	(1.336)	(1.080)	(0.104)	(2.135)	(1.788)	(0.161)	(1.547)	(1.408)	(0.140)
Average treatment effect (ATE)	7.210	5.891	0.113	6.405	4.349	0.298	7.757	6.906	-0.012
	(1.184)	(0.975)	(0.095)	(1.945)	(1.487)	(0.144)	(1.353)	(1.282)	(0.134)
ACME/ATE	0.143	0.444	0.584	0.280	0.530	0.446	0.097	0.457	-0.583

Notes: Estimates obtained following the procedure outlined in Appendix D of Imai et al. (2010a): we estimate the Linear Structural Equation Model using random effects regressions with question fixed effects as in Table 2, and we use the estimated sampling distributions to draw 100 simulations of potential mediators and potential outcomes. We average the differences of potential outcomes across the 100 simulations to obtain an estimate of the mediated effect. We repeat the procedure 500 times from bootstrap samples to obtain standard errors of the estimates.

Table C.2: Decomposition of post-debate treatment effect in mediated and direct effect

Sample:	Online		
Proposition alignment in:	Factual Beliefs	Confidence	Revealed Attitudes
Average causal effect mediated by s_i (ACME)	0.897	4.586	0.028
	(0.580)	(0.580)	(0.053)
Average direct effect (ADE)	5.177	3.735	0.047
	(1.565)	(1.303)	(0.143)
Average treatment effect (ATE)	6.074	8.320	0.074
	(1.538)	(1.229)	(0.134)
ACME/ATE	0.148	0.551	0.378

Notes: Estimates obtained following the procedure outlined in Appendix D of Imai et al. (2010a): we estimate the Linear Structural Equation Model using random effects regressions with question fixed effects as in Table 5, and we use the estimated sampling distributions to draw 100 simulations of potential mediators and potential outcomes. We average the differences of potential outcomes across the 100 simulations to obtain an estimate of the mediated effect. We repeat the procedure 500 times from bootstrap samples to obtain standard errors of the estimates.

D Materials

In this section we present instructions and questions of each survey, and discuss important features of implementation. We also explain any differences in the implementation of offline and online tournaments. All survey templates from the tournament in Munich (offline) and Amsterdam (online) are provided at https://osf.io/u7ekr/.

D.1 General remarks

We take several steps to collect high quality data in a confidential manner. First, all surveys begin with a cover page containing brief instructions to (i) inform subjects how much time they have to complete the survey, and (ii) remind subjects of the procedure to collect incentive compatible beliefs. Second, each survey is linked to the individual who filled it through a personal identifier. These IDs allow data to be collected and payments to be carried out confidentially. We ask debaters to enter their ID on the cover page of each of their surveys.

At offline tournaments, every study participant (debaters, judges, and enumerators) wears a name tag that includes their ID. Before collecting the survey, enumerators double-check that the ID entered by each debater on the cover page of their survey matches the one on the name tag. At online tournaments, participant IDs were sent to participants via email the day before the tournament. On the day of the tournament, before the first survey, each enumerator was responsible for double-checking

with a small group of debater and adjudicators that they had their participant IDs.

D.2 General instructions

We provided a document with general instructions for answering the surveys throughout the tournament. In particular, this explains how belief elicitations are incentivized using the Quadratic Scoring Rule for binarized outcomes (Harrison et al., 2014), how charitable allocations are paid out, and general payment procedures. All subjects are given 10 minutes to carefully read these general instructions right before the baseline survey begins. To make sure that procedures are adequately understood, if subjects miss their opportunity to read the general instructions we exclude them from the study.[3] The original content of these instructions is provided below. We separately report the instructions for offline and online tournaments.

D.2.1 General instructions (offline)

. . . . . .

General Instructions

Please read the following instructions carefully and keep them in mind, as they contain information that is relevant for the surveys we will ask you to complete during the next two days. We kindly ask you to use the time allocated to each survey to focus exclusively on answering the questions in front of you; throughout these times no information regarding the debates will be provided. Please answer each question carefully, don't use your phone and don't interact with others. Our instructions are never deceptive. All of your answers are treated confidentially and used for research purposes only.

Assessing factual statements

Spread across the various surveys, there are 34 questions that are marked by an "$", for which you can earn money. After you completed the last survey, we will pay you based on one randomly selected answer. While you will get paid for only one of your answers, every question might be the one that counts.

[3] They are allowed to answer the surveys, but their data is discarded.

Questions marked by an "$" ask you to state the likelihood (in percent) that a given statement is true. Most such statements are designed to assess your factual knowledge. There will be no trick questions. Moreover, all sources we refer to actually exist and are of high quality, but the actual fact may be either true or not true. As an example, consider the following statement.

According to Eurostat, more than 30 percent of live births in Germany in 2016 were outside of marriage.

This statement is true if Eurostat indeed reported this finding. It is false if Eurostat reported a different finding. You will be asked to provide your belief as to how likely you think it is that this statement is true. If this answer is selected for payment, you will earn either 30 euros or nothing. The procedure that determines how likely it is that you win the 30 euros assures that the closer you are to the correct answer (either 0 or 100 percent), the higher is your probability of winning the money.

Moreover, the procedure assures that you maximize your chance of winning money by stating your true belief (between 0 and 100 percent). So if you are almost certain that a given statement is true, then you should state a belief that is very high. If you are almost certain that a given statement is false, then you should state a belief that is very low. If you are completely uncertain, you maximize your chance of winning by stating a belief that is close to 50 percent.

The Procedure Box below provides more comprehensive information about the exact payment mechanism. But note that it is not important that you understand the procedure in detail. What matters is that you know that you maximize your probability of winning when you report your true belief - if you under- or overstate your belief, you will reduce your chance of winning the 30 euros.

Donating to charities

For some questions in the survey, you will be able to allocate monetary endowments between different charities. This is money that we make available from our budget for you to allocate, according to your preferences, to charities that have different missions. One of the allocations you make will be selected at random and we will transfer the money to the relevant charities. While we will implement only one of your allocations, every allocation might be the one that counts.

The surveys will also feature further questions that allow you to earn more money for yourself. The instructions for these questions are simple and will be provided above the relevant question.

Procedure Box

How a given answer maps into your chance of winning 30 euros is based on a formula. This formula is designed to make sure that you maximize your chance of winning if you report your true belief that a given statement is true.

Suppose that the correct answer is given by R, which is equal to 1 if the statement is true and 0 if the statement is false. The variable r is your report–the likelihood that you attribute to the statement being true (from 0 to 100 percent). The winning probability for the prize is then given by:

$$\text{winning probability} = 100 - 100 \times (R - r/100)^2$$

Example: Suppose again that you are tasked with assessing the following statement: *According to Eurostat, more than 30 percent of live births in Germany in 2016 were outside of marriage*. And suppose that your belief that the statement is true is 63 percent. The following table shows your winning probability based on the formula. The columns represent a number of hypothetical answers you may give. As you can see, you maximize your chance of winning by reporting your true belief.

	Report 1	Report 2	**Report 3**	Report 4
Hypothetical report	22	35	**63**	89
Expected winning probability if **your belief** that the statement is true is **63%**	59.9%	68.9%	**76.7%**	69.9%

Payment

On Sunday, we will pay out your earnings in cash. To determine your earnings for the assessment of factual statements, we first randomly draw the question that is relevant for your payment. We then determine your winning probability based on the true answer and your reported answer. Finally, a computer program constructs a virtual urn with only white and black balls, where the share of white balls equals your winning probability. If the computer then draws a white ball from the urn, then you will win the 30-euro prize. This is a fair and transparent procedure to pay you the prize with the winning probability you have earned based on the quality of your answers.

If the question that is drawn for payment is from a round that you missed, then there will be no new draw and you will not earn any money for this type of question. If you would like us to send you receipts of the charity donation based on your choice, then please leave us your email address when

you collect your payment.

. . . . . .

D.2.2 General instructions (online)

Below we present general instructions for debaters. Adjudicators had similar instructions.

. . . . . .

General Instructions

Please read the following instructions carefully and keep them in mind, as they contain information that is relevant for the surveys we will ask you to complete during the next two days. We kindly ask you to use the time allocated to each survey to focus exclusively on answering the questions in front of you; throughout these times no information regarding the debates will be provided. Please answer each question carefully, don't use your phone and don't interact with others. Our instructions are never deceptive. All of your answers are treated confidentially and used for research purposes only. If anything is unclear, don't hesitate to contact the researchers (egon.tripodi@essex.ac.uk) and we will aim to reply immediately.

Assessing factual statements

Spread across the various surveys, there are 35 questions that are marked either by a "(5 euro prize)" or by a "(50 euro prize)", for which you can earn money. After you completed the last survey, we will pay you based on one randomly selected answer among those marked by a "(5 euro prize)" and one randomly selected answer among those marked by a "(50 euro prize)". While you will get paid for only two of your answers, every question might be the one that counts.

Questions marked by a "(5 euro prize)" or "(50 euro prize)" ask you to state the likelihood (in percent) that a given statement is true. Most such statements are designed to assess your factual knowledge or intuition. There will be no trick questions. Moreover, all sources we refer to actually exist and are of high quality, but the actual fact may be either true or not true. As an example, consider the following statement.

According to Eurostat, more than 30 percent of live births in Germany in 2016 were outside of marriage.

This statement is true if Eurostat indeed reported this finding. It is false if Eurostat reported a different finding.

You will be asked to provide your belief as to how likely you think it is that this statement is true. You will state a probability between 0 and 100 percent.

If this answer is selected for payment, you will earn either monetary prize or nothing. The procedure that determines how likely it is that you win the monetary prize assures that the closer you are to the correct answer (either 0 or 100 percent), the higher is your probability of winning the money. The monetary prize is 5 euro for questions marked by a "(5 euro prize)", and 50 euro for questions marked by a "(50 euro prize)".

Moreover, the procedure assures that you maximize your chance of winning money by stating your true belief (between 0 and 100 percent). So if you are almost certain that a given statement is true, then you should state a belief that is very high. If you are almost certain that a given statement is false, then you should state a belief that is very low. If you are completely uncertain, you maximize your chance of winning by stating a belief that is close to 50 percent.

The Procedure Box below provides more comprehensive information about the exact payment mechanism. But note that it is not important that you understand the procedure in detail. What matters is that you know that you maximize your probability of winning when you report your true belief - if you under- or overstate your belief, you will reduce your chance of winning the monetary prizes.

Procedure Box

How a given answer maps into your chance of winning a monetary prize is based on a formula. This formula is designed to make sure that you maximize your chance of winning if you report your true belief that a given statement is true. You may click the following button to get more explanation on the exact procedure.

[More info][4]

Donating to charities

For some questions in the survey, you will be able to allocate monetary endowments between different charities. This is money that we make available from our budget for you to allocate, according to your preferences, to charities that have different missions. One of the allocations you make will be selected at random and we will transfer the money to the relevant charities. While we will implement only one of your allocations, every allocation might be the one that counts.

Payment

After the tournament, we will pay out your earnings using the payment data you provided. To deter-

mine your earnings for the assessment of factual statements, we first randomly draw the question that is relevant for your payment. We then determine your winning probability based on the true answer and your reported answer. Finally, a computer program constructs a virtual urn with only white and black balls, where the share of white balls equals your winning probability. If the computer then draws a white ball from the urn, then you will win the monetary prize. This is a fair and transparent procedure to pay you the prize with the winning probability you have earned based on the quality of your answers. We repeat this procedure for the 5 euro prize and for the 50 euro prize. You may win none of the prizes, either, or both.

If the questions drawn for payment are from rounds that you missed, then there will be no new draw and you will not earn any money for this type of question. If you would like us to send you receipts of the charity donation based on your choice, then please email us at egon.tripodi@essex.ac.uk including your participant ID (D###).

. . . . . .

D.3 Baseline survey

A 25-minute baseline survey includes the following items:

- Age (open field, suggested to provide a numeric answer).
- Gender (open field).
- Nationality (open field).
- Political ideology scale: "*In politics people sometimes talk of "left" and "right". Where would you place yourself on this scale, where 0 means the left and 10 means the right?*" (check box).
- Years actively debating on a regular basis. Options: "Less than a year", "1 to 2 years", "3 to 4 years", "At least 5 years". (check box)
- Times debater got to semifinals in Open/IV tournaments (open field).
- "*What do you think makes a good debater*". Options: "Choosing arguments strategically", "Confidence in own position", "Debating experience", "Factual knowledge", "Eloquence" (ranking).[5]
- Incentivized belief elicitation on fifteen factual statements: for each such statement subjects state how likely it is that the fact is true (open field, suggested to provide a numeric answer from 0 to 100).

[5]Dropped at online tournaments.

- "*Did you take part as a speaker at the Munich Research Open 2019?*". Options: "Yes", "No" (check box).[6]

A key component of this survey was to gather beliefs at baseline regarding the motions that subjects were going to debate. At the same time, we had to avoid that our questions revealed the motions. To obfuscate the relation of these belief elicitations and the motions we elicit beliefs over whether 15 factual statements are true: 5 such statements relate to the in-round motions, 7 are decoy questions, and 3 are control questions.[7] For each team of debaters, control questions are drawn from a pool of 6 questions, and the questions that were not selected for the baseline survey are then included in the endline survey. Comparing responses to the control questions at baseline and endline by different debaters helps uncover to what extent debaters research the questions. We find no evidence that they do.

Decoy questions are designed to look like they could relate to plausible motions for debate. Control questions are facts that not necessarily relate to typical debate topics.

For each motion, we devise multiple factual statements that we phrase as binary states to capture alignment of beliefs with the persuasion goal. Any given question may not have a tight enough link to the motion in debaters' minds or give rise to a high degree of certainty in debaters' beliefs and may therefore be ill-suited to pick up a treatment effect. To diversify this risk, at offline tournaments, we come up with 4 questions (A, B, C, D) for each motion and administer them as illustrated in the table below: at baseline, debaters are asked either about fact A or B; pre-debate, debaters are asked either about fact D or C; post-debate debaters are asked either about fact B and C or A and D.

This approach also ensures that (i) no debater is asked the same question twice, and (ii) we protect the baseline and pre-debate belief elicitations from any potential information spillovers.

Timing:	Beginning of Day 1	Day 1 or Day 2	
	Baseline	pre-debate	post-debate
Subgroup 1	A	D	B, C
Subgroup 2	B	C	A, D

After observing very strong compliance at offline tournaments with our request to not discuss the

[6] Only in Rotterdam.

[7] In Rotterdam, 4 statements relate to the in-round motions, and 8 are decoy questions. At the online tournaments, we have only 12 factual statements because we drop control questions.

questions of the survey, we decided to reduce the number of factual statements per motion to three and randomize the order as follows. This approach still ensures that no debater is asked the same question twice.

Timing:	Beginning of Day 1	Day 1 or Day 2	
	Baseline	pre-debate	post-debate
Subgroup 1	A	B	C
Subgroup 2	B	C	A
Subgroup 3	C	A	B

D.4 Pre-debate survey

This 5 minute survey is handed out before each debate begins and after the preparation time. It includes:

- Incentivized belief elicitation on two factual statements: for each such statement subjects state how likely it is that the fact is true (open field, suggested to provide a numeric answer from 0 to 100).
- Choice of one of 9 monetary allocations, along a concave budget, between a baseline charity and a charity aligned with one of the sides represented in the debate. For an illustration see Figure D.1.
- Questions on the number of arguments considered during preparation time in favor of the proposition:
  - i *How many good arguments did you come up with during the preparation time in favor of the proposition?* (open field, suggested to provide a numeric answer)
  - ii *How many of these arguments would you consider to be very strong?* (open field, suggested to provide a numeric answer between zero and the answer to the previous question)
- Questions on the number of arguments considered during preparation time against the proposition:
  - i *How many good arguments did you come up with during the preparation time against the proposition?* (open field, suggested to provide a numeric answer).
  - ii *How many of these arguments would you consider to be very strong?* (open field, suggested to provide a numeric answer between zero and the answer to the previous question).

Below you see nine potential ways in which you could allocate charitable donations—that are paid by us on your behalf—between two charitable organizations: Oxfam and The Planetary Society .

Oxfam is a major nonprofit group with an extensive collection of operations. Oxfam's programs address the structural causes of poverty and related injustice and work primarily through local accountable organizations, seeking to enhance their effectiveness

The Planetary Society is the world's largest and most influential non-profit space organization. The society advocates for space and planetary science funding in government, invests in inspiring educational programs, and funds groundbreaking space science and technology

How would you like to allocate these donations? (check only one box)

Choose one option	☐	☐	☐	☐	☐	☐	☐	☐	☐	
You want to give	0.0	1.3	2.5	3.7	5.0	5.9	6.4	6.7	7.0	euro to Oxfam
and	7.0	6.7	6.4	5.9	5.0	3.7	2.5	1.3	0.0	euro to the Planetary Society
A total of	7.0	8.0	8.9	9.6	10.0	9.6	8.9	8.0	7.0	*euro goes to charity*

Figure D.1: Illustration of charitable donations allocation question

Both factual statements are meant to capture whether beliefs are aligned with the motion after the debate. The first statement features a real-world fact. The second statement elicits confidence in the arguments of the proposition side of the debate by asking:

Statement: Excluding the debate happening in this room, in at least half of the parallel debates of this round, one of the two teams on the Government side of this motion will rank 1st.

Q2$: How likely do you think it is that the above statement is true? ___% (write a number from 0 to 100)

For each motion, we select two charities that we expect to be aligned with one side of the debate. We randomly determine which of these two charities features in the pre-debate survey. The other charity features in the post-debate survey. In Munich we also randomize the baseline charity between Oxfam and Opportunity International. At all other tournaments, the baseline charity is always Opportunity International.

D.5 Post-debate survey

This 5 minute survey is handed out right after each debate. It includes:

- Incentivized belief elicitation on two factual statements: for each such statement subjects state how likely it is that the fact is true (open field, suggested to provide a numeric answer from 0 to 100). The first is for the factual belief elicitation and the second is for the confidence

elicitation.[8]

- Subjective ranking of team performance in the debate.[5]
- Choice of one of 9 monetary allocations, along a concave budget, between a baseline charity and a charity aligned with one of the sides represented in the debate. For an illustration see Figure D.1.
- Questions on the number of arguments considered during preparation time in favor of the proposition:[9]

    i *After taking part in this debate, how many good argument can you think of in favor of the proposition?* (open field, suggested to provide a numeric answer)

    ii *How many of these arguments would you consider to be very strong?* (open field, suggested to provide a numeric answer between zero and the answer to the previous question)
- Questions on the number of arguments considered during preparation time against the proposition:[9]

    i *After taking part in this debate, how many good argument can you think of against the proposition?* (open field, suggested to provide a numeric answer).

    ii *How many of these arguments would you consider to be very strong?* (open field, suggested to provide a numeric answer between zero and the answer to the previous question).

D.6 Endline survey

At offline tournaments the endline is a lengthier survey in which we measure perceived alignment of each factual statement and charity with respect to the motion along with a short exit set of questions. Due to time constraints of the online format we drop the measurement of perceived alignment, and we bundle this survey with the last post-debate survey.

D.7 Endline survey (offline tournaments)

This 20-minute survey takes place right after the last round of debates. It includes:

[8] At offline tournaments, both statements are for the factual belief elicitation, and we do not elicit confidence post-debate.

[9] Added at online tournaments.

- A question that we use to assess how debaters think that beliefs about facts that we ask and charities they can donate to relate to alignment with the motions. An illustration of the precise wording of this question is provided in Figure D.2.
- Incentivized belief elicitation on four factual statements: for each such statement subjects state how likely it is that the fact is true (open field, suggested to provide a numeric answer from 0 to 100).
- Open text box in which subjects are asked to tell us what they think the research was about.[10]

Three of the four factual statements are control questions of the kind included in the baseline survey. One fact pertains the performance of two actual debaters in the Munich Research Open, and had a longer preamble than other belief elicitation questions:

The next question is about the performance of two actual debaters in a different tournament: the Munich Research Open that took place two weeks ago[11]*. We will call them debater A and debater B. Both debaters were representing the Government in the motion that "THBT governments should stop funding scientific programmes that have no immediate benefit for humankind (such as space travel and exploration, human cloning)", but they gave different responses to the factual question in the pre-debate survey:*

Debater A believed that the statement "More than 10 of the following 15 innovations are a consequence of inventions made in the pursuit of space travel: camera phones, scratch resistant lenses, electric light, CAT scans, LEDs, land mine removal, athletic shoes, penicillin, water purification systems, the internet, home insulations, wireless headsets, baby formula, portable computers" was true with 75% chance. Debater B believed that the same statement was true with 10% chance.

We asked judges to provide a broad measure of each debaters' persuasiveness. Now consider the following statement.

Statement: Debater A obtained a higher persuasiveness score than Debater B in the relevant debate.

Q6$: How likely do you think it is that the above statement is true? ___ % (write a number from 0 to 100)

[10] We felt that the alignment question was revealing too much of what the study was about, so to get a better sense of whether subjects understood what hypotheses were being tested with the data collected in pre-debate and post-debate surveys, in Rotterdam, we decided to move this question to the last post-debate survey.

[11] In Rotterdam. In Munich, the orange text is replaced by "this tournament".

Q1: For this question, you can earn up to 5 euros. Consider the following motion, which was debated during the event:

"During periods of national housing shortages, this House would forcibly take ownership of privately owned homes which are not lived in by their owners".

Now consider someone who is strongly in favor of this motion, i.e. someone whose personal views are strongly aligned with the motion. For each statement in the table below, please indicate whether such a person, who is strongly aligned with the motion, is more likely to believe that the statement is true or more likely to believe that the statement is false. For each statement (each row of the table), give your answer by entering one (and only one) cross in the appropriate box.

We will randomly select one of these six statements and pay you based on your selection as follows: You will earn 5 euros for sure if your response is the same as the response that is selected most frequently by all other participants answering the same question.

	Someone aligned with the motion is…		
	… more likely to believe that this statement is **true**	… more likely to believe that this statement is **false**	.. equally likely to believe that this statement is **true or false**
Statement 1: According to the English Housing Survey, the number of second homes in the UK more than doubled between 1995 and 2013			
Statement 2: Under current UK regulation, squatters who live in and maintain unoccupied buildings enjoy protection under the law and can never be evicted without a court order			
Statement 3: According to an academic study published this year, over 5 percent of properties in England and Wales are low-use properties, defined as a property that is not registered as the primary residence of any individual			
Statement 4: According to research by the newspaper the Independent in 2018, more than one third of new-build luxury apartments and houses in Central London lies empty			
Statement 5: Action on Empty Homes* is an NGO supporting a cause that is especially important.			
Statement 6: The Land Is Ours** is an NGO supporting a cause that is especially important.			

*Action on Empty Homes is a UK NGO campaigning for more empty homes to be brought into use for people in housing need. It raises awareness of the waste of long-term empty homes and campaigns for changes to national policy to bring more homes into use..
**The Land Is Ours campaigns peacefully for access to the land, its resources, and the decision-making processes affecting them. Among other things, it advocates 'Use It Or Lose It' programme where empty buildings are forfeit or put on a tax escalator, where the owner can lose title after one year.

Figure D.2: Example of aligment question in the endline survey

D.7.1 Endline survey (online tournaments)

This 5-minute survey is tagged to the last post-debate survey. It includes:

- Incentivized belief elicitation on five factual statements: for each such statement subjects state how likely it is that the fact is true (open field, suggested to provide a numeric answer from 0 to 100). The five statements mirror the ones used at Baseline to measure alignment of factual beliefs with the proposition of rounds 1-5.
- Open text box in which subjects are asked to tell us what they think the research was about.

- Questions on knowledge of the researchers and their research.

D.8 Adjudicator survey

The adjudicator survey differs substantially between offline and online tournaments. Hence, we present it separately.

D.8.1 Adjudicator survey (offline tournaments)

Adjudicators are asked to independently provide individual scores of each debater's overall persuasiveness before filling out the shared score sheet with other adjudicators.

Adjudicators are asked to provide a broad persuasiveness score, on a scale from 1 to 10 where 1 is "Not at all persuasive" and 10 "Extremely persuasive". The original instructions given to judges on how to answer and interpret this question are provided below:

Without discussing with the other judges, please evaluate the persuasiveness of each debater. We consider a debater persuasive, if she would do well at convincing a general audience of her position. Therefore, please provide a broad measure of persuasiveness that captures the quality of arguments as well as speaking ability, body language and any other attribute that makes a speech persuasive to a general audience.

To ensure that the adjudicators provided independent persuasiveness scores, we asked them to fill out these surveys during the debate. Adjudicators on the panel painstakingly take notes of each speech and generally do not interact with each other during the debate. We collected the surveys before any deliberation of the panel took place.

D.8.2 Adjudicator survey (online tournaments)

This 5-minute survey is collected at the same time as the pre-debate survey among debaters. It includes:

- Incentivized belief elicitation on three factual statements: for each such statement subjects state how likely it is that the fact is true (numeric answer from 0 to 100). These are three factual statements of the round.
- For each of these factual statements as well as for the confidence question, we ask adjudicators to also predict the average response of debaters on the proposition side and on the opposition side of the debate (numeric answer from 0 to 100).

- For each of the two motion related charities of the round, we elicit an hypothetical choice of one of 9 monetary allocations, along a concave budget, between the baseline charity and the charity aligned with one of the sides represented in the debate.
- For each of the two motion related charities of the round, we also ask adjudicators to predict the average choice of debaters on the proposition side and on the opposition side of the debate.

D.9 Enumerator survey

The enumerator in the offline tournaments filled out a survey during the debate that includes the following items:

- A count of the times not speaking debaters try to interrupt the speaker (through Points of Information).
- A subjective rating of how heated each debaters' argumentation is coming across (on a scale from 1 to 5).[12]
- For each of the four facts related to the motion over which we elicit debaters beliefs, and for both the motion related charities, note whether these were mentioned during the debate.

We drop this survey at the online tournaments.

D.10 Ballot

Adjudicators fill our the ballot: The official form that determines the outcomes of the debating tournament. This form includes:

- Name and position of each team in the debate
- Ranking of the four teams in the debate (from First to Fourth, with no possibility for ties)
- Individual speaker scores (on a scale from 50 to 100)

After a debate is over, speakers leave the room to let adjudicators on the panel privately discuss the performance of each debater. This discussion takes approximately 15 minutes during which the

[12]Enumerators were instructed to write down this score for each debater at the end of the speech. They could however revise this score for debaters that acted particularly heatedly during other debaters' speeches.

arguments presented by each debater are technically analyzed. A technical analysis is particularly relevant to the assignment of individual speaker scores, which are supposed to be assigned an a objective scale that applies to any British Parliamentary performance.[13] The ballot is filled out at the end of this discussion.

[13] An example of such scale can be found at https://debate.uvm.edu.

Table D.1: Alignment of facts with motions

Factual statements, organized by motions[14]	Alignment
This House believes that governments should stop funding scientific programs that have no immediate benefit for humankind (such as space travel and exploration, human cloning)	
1. The European Space Agency's annual budget exceeds 4% of the EU budget	Proposition
2. More than six Europeans out of ten agree that space technologies have a role to play in avoiding threats like asteroids, comets, and space debris collisions	Opposition
3. More than 10 of the following 15 innovations are a consequence of inventions made in the pursuit of space travel: camera phones, scratch resistant lenses, electric light, CAT scans, LEDs, land mine removal, athletic shoes, penicillin, water purification systems, the internet, home insulation, wireless headsets, baby formula, portable computers	Opposition
4. A study in a leading economics journal in 1998 finds that the majority of R&D spending by the US government goes into wages for scientists, which in turn does little to increase the number of scientists	Proposition
This House believes that Western States should permanently revoke the citizenship of citizens who join terrorist organisations	
1. From 2015 to 2017 there were more than 50 separate Islamic terrorist attacks in the EU	Proposition
2. According to the UN's Basic Human Right's Reference Guide on the right to a fair trial and due process in the context of countering terrorism, the citizenship of people suspected or proven of having been part of terrorist organisations may never be revoked	Opposition
3. In Germany, the law governing citizenship already permits to strip those with dual citizenship of their German citizenship if they join a foreign army	Proposition
4. The two main perpetrators of the 2015 attacks on the offices of the satirical newspaper Charlie Hebdo had fought with IS in Syria prior to the attack	Proposition
This House regrets the EU's introduction of freedom of movement	
1. In a much-cited academic article from 2012, researchers from University College London found that immigration increased wages in the UK, both at the bottom and at the top of the income distribution	Opposition
2. More than 35% of UK citizens interviewed for the Eurobarometer in 2018 think that the Schengen Area has more disadvantages than advantages for the UK	Proposition
3. According to a 2018 paper by researchers from the University of Munich, emigration within Europe positively contributes to innovation in source countries, i.e. countries people emigrate from	Opposition
4. Less than half of Europeans agree that integration of immigrants has been a success in their local area, city or country	Proposition
This House would suspend trade union powers and significantly relax labour protection laws in times of economic crisis	
1. A 2015 survey by the Employment Policies Institute shows that a majority of economists thinks that a USD 15 minimum wage will reduce the number of jobs	Proposition
2. According to a study of 21 Eastern European economies published in 2017, members of labor unions are less likely to lose their job during an economic crisis	Opposition
3. The fraction of the U.S. population that approves of labor unions dropped by more than one third from the mid-50s to 2009, according to the Gallup poll	Proposition
4. In a 2005 study of OECD countries, economists from Harvard University and the University of Bonn find that greater labor market flexibility (e.g. due to weaker trade unions) is associated with greater employment	Proposition
This House believes that causing deliberate harms to enemy civilians, by the weaker side, is a justified tactic in asymmetrical warfare	
1. Research on the psychological effect of violence against Israeli civilians shows that such violence caused a hardening of attitudes, stronger opposition to political reconciliation with perpetrators, and an increase in support for counter-terrorist measures	Proposition
2. Nelson Mandela, who was awarded the Nobel Peace Prize for facilitating South Africa's peaceful transition into democracy, was also the co-founder of the violent paramilitary wing of the African National Congress and was classified as a terrorist by the US until 2008	Proposition
3. Willingness to resort to violence was an integral part of the Birmingham campaign, which is widely credited with bringing about civil rights and desegregation in the United States of the 1960s	Proposition
4. Studies in political science consistently find that that rebel groups that use indiscriminate violence against civilians are more likely to achieve political goals	Proposition
During periods of national housing shortages, this House would forcibly take ownership of privately owned homes which are not lived in by their owners)	
1. According to the English Housing Survey, the number of second homes in the UK more than doubled between 1995 and 2013	Proposition
2. Under current UK regulation, squatters who live in and maintain unoccupied buildings enjoy protection under the law and can never be evicted without a court order	Opposition

[14]We redacted one motion from this table due to its sensitive content. Our analysis is based on all 19 motions.

3. According to an academic study published this year, over 5 percent of properties in England and Wales are low-use properties, defined as a property that is not registered as the primary residence of any individual — Proposition

4. According to research by the newspaper the Independent in 2018, more than one third of new-build luxury apartments and houses in Central London lies empty — Proposition

This House believes that states should aggressively fund geoengineering projects instead of attempting to mitigate the effect of climate change

1. Germany's experience with renewable energy promotion (i.e. its Renewable Energy Sources Act (EEG)) is often used as a model to be replicated elsewhere. Instead, a widely cited scientific study from 2010 argues that the German government's support of renewables has resulted in massive expenditures (annual feed-in tariffs of over 7 billion euros) that show little long-term promise for stimulating the economy, protecting the environment, or increasing energy security — Proposition

2. According to recent data from the Climate Action Tracker, more than one third of the surveyed countries are well on track to meet the CO2 emission targets they imposed on themselves under the Paris agreement — Opposition

3. Even the US, which has not supported recent global efforts to fight climate change by means of reducing CO2 emissions, has been enthusiastic in its support for geoengineering projects, as evidenced by its support for the U.N. resolution on geoengineering — Proposition

4. A 2018 study by two prominent economists from MIT argues that increased investments in geoengineering may also increase efforts to improve clean energy technologies — Proposition

This House regrets the decision to let the FARC (i.e. The Revolutionary Armed Forces of Colombia - People's Army) run as a political party.

1. Shortly after the 2016 peace deal with FARC, Colombia has been experiencing a resurgence of violence. The number of homicides is up by more than 7% in 2018 compared to the previous year — Proposition

2. In 2016, the Nobel peace prize was jointly awarded to Colombian president Santos and the leader of FARC, Rodrigo Londoño, for their "resolute efforts to bring the country's more than 50-year-long civil war to an end" — Opposition

3. In March 2017, the Colombian government reported that more than 25% of the estimated 6'900 FARC fighters refused to disarm — Proposition

4. Towards the end of the peace deal negotiations between the Colombian government and FARC, NGOs like Amnesty International and Human Rights Watch as well as the Colombian Conservative party criticized the peace deal for being too lenient on perpetrators of human rights violations — Proposition

When tech companies own platform utilities and platform products, this House would break them up.

1. According to a 2018 survey from the Pew Research Center, over 50% of Americans believe that major tech companies have too much power and influence in today's economy — Proposition

2. The UK government's digital competition expert panel, chaired by Professor Furman who was chief economic advisor in Obama's presidency, issued a report just two weeks ago rejecting the widely held view that "digital platforms are natural monopolies where only a small number of firms can succeed" — Opposition

3. According to a 2018 survey from the Pew Research Center, over 60% of Americans believe that major tech companies should be more regulated than they currently are — Proposition

4. A 2018 survey of 1200 sellers on the Amazon platform, conducted by the independent market research firm Feedvisor, finds that over 40% of private sellers on Amazon fear that the company will take away their seller privileges and over 60% of them fear Amazon competing directly with them — Proposition

This House believes that it is in the interest of men for gender roles (e.g. work-life balance, choice of occupation, assertiveness expectations) to be dismantled

1. A study by the London School of Economics finds that traditional gender identity deters men from entering female-dominated professions. — Proposition

2. A study of over seventy thousand individuals across several countries finds that greater gender equality is associated with better mental health for men. — Proposition

3. A paper in the journal Labor Economics documents that men in male-dominated professions are happier with their jobs than men in female dominated professions. — Opposition

This House believes that humanitarian organizations should primarily use apolitical aid as opposed to politicized action

1. A survey in Iraq in 2006/7 showed that the perceived neutrality and impartiality of humanitarian aid was reduced by collaborations between humanitarian and military or political actors. — Proposition

2. A 2012 Ipsos poll in the UK found that "political bias" was the most frequently cited reason among those who reported a decrease in trust in charities. (FALSE - 5th important reason). — Proposition

3. Political activism by humanitarian organisations across the world was key in ending apartheid in South Africa, since both Margaret Thatcher and Ronald Reagan cited said activism in their decision to sanction the apartheid regime. — Opposition

This House would allow people to sign income share agreements

1. According to a study by Purdue University published this year, there is a large negative effect of income share agreements on subsequent student grade. — Opposition

2. According to a study by the Institute for College Access and Success, a move away from student debt to income share agreements would increase the share of low-income students in higher education, because low-income students tend to prefer income share agreements. — Proposition

3. In 2019, U.S. senator Elisabeth Warren sent a letter to secretary of education Betsy DeVos, arguing that income share agreements include some of the most exploitative terms in the private student loan industry.	Opposition
This House believes that developing nations should prioritise investment in future technologies (e.g. 5G, cybersecurity, green energy) over traditional sectors (e.g. agriculture, manufacturing)	
1. Renewable energy will account for 80% of total global energy jobs by 2050 according to a 2020 study in the Journal of Technological Forecasting and Social Change.	Proposition
2. Over 60 percent of developing nation populations live in rural areas, often far away from urban hubs and with little access to high-quality education.	Opposition
3. According to a paper published in the journal Renewable and Sustainable Energy Reviews, the introduction of smart grid technology in China, India and Brazil has been an economic success story.	Proposition
This House supports engaging private military companies to combat terrorism	
1. According to an article in the Journal of Peace Research, private military companies have a destabilizing effect by creating their own demand for further conflict.	Opposition
2. This year, the U.N. accused a private military company in Libya of using illegal chemical weapons	Opposition
3. According to a paper published in the journal of African Studies, private military companies were an integral part of the successful military campaign against the terrorist group Boko Haram in Nigeria	Proposition
This House regrets modern medicine's prioritisation of longevity over quality of life (e.g. extending lifespan vs. palliative care).	
1. A 2017 review of clinical trials on the effect of palliative care on patients' self-reported quality of life concluded that there was no such effect.	Opposition
2. A review of 38 academic studies concluded that more than 50% of patients near the end of life receive completely ineffective medical interventions aimed at extending their life.	Proposition
3. In a 2015 article, researchers from the University of Cambridge argue that the focus of medical research on increasing longevity, rather than the quality of life, is damaging to patients' health and the economy.	Proposition
This House proposes a world where, after COVID-19 is no longer a threat to public health and safety, remote work remains the norm	
1. According to the U.S. Census Bureau, the average American worker spent over 15 full days commuting in 2018.	Proposition
2. A 2017 report by the International Labor Organization found that remote workers report better work-life balance and higher productivity.	Proposition
3. According to a November 2020 survey by consulting agency PriceWaterhouseCoopers, less than 50% of employers say that their shift to remote work has been a success.	Opposition
This House opposes political consumerism	
1. An academic assessment of the impact of 21 consumer boycott announcements found that they had no effect on the market value of the target firms.	Proposition
2. According to Helen Lewis, a prominent feminist, political consumerism tends to bring about superficial reactions by companies rather than genuine reform.	Proposition
3. Research in the Journal of Civil Society shows that engaging in political consumerism has the benefit of also leading to a higher level of engagement in conventional political activism.	Opposition
This House supports the rise of private cities in the developing world	
1. A 2017 article in the Journal of Urban Geography criticizes China for using private cities as a geopolitical tool for its expansionist agenda.	Opposition
2. Lavasa, one of the first private cities in India, has been largely abandoned by residents as its corporate owners ran out of funds for maintenance.	Opposition
3. A 2020 article, published in the Journal of Development Economics, documents how a UN-funded private city in Ghana significantly improved the health and educational outcomes of it's inhabitants.	Proposition

Table D.2: Alignment of charitable causes with motions

Charitable causes, organized by motions[15]	Alignment
This House believes that governments should stop funding scientific programs that have no immediate benefit for humankind (such as space travel and exploration, human cloning)	
1. **The International Space University** develops the future leaders of the world space community. It encourages the innovative development of space for peaceful purposes: to improve life on Earth and advance humanity into space	Opposition
2. **The Planetary Society** is the world's largest and most influential non-profit space organization. The society advocates for space and planetary science funding in government, invests in inspiring educational programs, and funds groundbreaking space science and technology	Opposition
This House believes that Western States should permanently revoke the citizenship of citizens who join terrorist organisations	
1. **The Active Change Foundation** is based in the UK and provides a holistic approach to neutralising extremism and violence on both an individual and community level. Its chief executive is an outspoken critic of those actors within the UK that favor stripping individuals of their citizenship for being involved with terrorist organisations	Opposition
2. **Human Rights Watch** defends the rights of people worldwide. It scrupulously investigates abuses, exposes the facts widely, and pressures those with power to respect rights and secure justice. It has been a vocal defender of the right to citizenship for all people	Opposition
This House regrets the EU's introduction of freedom of movement	
1.**The European Movement UK** is a grass-roots, independent, pro-European organisation. One of its main goals is to safeguard the freedom of movement made possible by membership of the EU, both for UK citizens who want to travel and work abroad and for citizens of other EU countries who want to come to the UK to work and to live	Opposition
2. **ACT4FreeMovement** stands for Advocacy, Complaints, Trainings for Freedom of Movement. The organization campaigns for freedom of movement with EU citizens. The goal is to increase the capacity of EU citizens to effectively secure access to and knowledge of their rights, as well as build public awareness and political support for mobile citizen rights	Opposition
This House would suspend trade union powers and significantly relax labour protection laws in times of economic crisis	
1. **The European Trade Union Confederation** speaks with a single voice on behalf of European workers to have a stronger say in EU decision-making. It aims to ensure that the EU is not just an economic union but also a Social Europe, where improving the well-being of workers and their families is an equally important priority	Opposition
2. **The Living Wage Foundation** is a campaigning organization in the United Kingdom, which aims to persuade employers to pay a Living Wage, an independently calculated and recommended minimum wage to cover workers' basic needs	Opposition
This House believes that causing deliberate harms to enemy civilians, by the weaker side, is a justified tactic in asymmetrical warfare	
1. **The Israel Trauma Center** for Victims of Terror and War is an apolitical organization providing multidisciplinary treatment and support to direct and indirect victims of trauma due to terror and war in Israel	Proposition
2. **Muslim Aid** is an Islamic Charity, which has been actively working in Gaza since 2006. It helps vulnerable people to obtain essentials like food and medical supplies, which are scarce as importing and exporting has been made difficult	Proposition
During periods of national housing shortages, this House would forcibly take ownership of privately owned homes which are not lived in by their owners)	
1. **Action on Empty Homes** is a UK NGO campaigning for more empty homes to be brought into use for people in housing need. It raises awareness of the waste of long-term empty homes and campaigns for changes to national policy to bring more homes into use.	Proposition
2. **The Land Is Ours** campaigns peacefully for access to the land, its resources, and the decision-making processes affecting them. Among other things, it advocates 'Use It Or Lose It' programme where empty buildings are forfeit or put on a tax escalator, where the owner can lose title after one year	Proposition
This House believes that states should aggressively fund geoengineering projects instead of attempting to mitigate the effect of climate change	
1. **Geoengineering Monitor** aims to be a timely source for information and critical perspectives on climate engineering. The goal is to serve as a resource for people around the world who are opposing climate geoengineering and fighting to address the root causes of climate change instead	Opposition
2. **The Environmental Defense Fund** addresses today's most urgent environmental challenges by focusing on the solutions that will have the biggest impact, such as removing obsolete rules that hamper the clean energy market in the U.S. It favors a strategy of reducing CO2 emissions over geoengineering	Opposition
This House regrets the decision to let the FARC (i.e The Revolutionary Armed Forces of Colombia - People's Army) run as a political party.	

[15]We redacted one motion from this table due to its sensitive content. Our analysis is based on all 19 motions.

1. **Justice for Colombia** is a British NGO whose primary goal is to give a political voice internationally to Colombian civil society. It has been campaigning to help Jesús Santrich, a lead FARC negotiator of the peace deal who was going to take a seat into parliament in 2018, get justice. The US incarcerated him without providing any evidence of Santrich's crime to the Colombian government — Opposition

2. **Strangers to Peace** is a documentary project of film maker Noah DeBonis which follows the life of ex-FARC guerrillas during their reintegration process. If funded, the film aims to enrich viewer's understanding of a marginalized community through tales of personal and social redemption — Opposition

When tech companies own platform utilities and platform products, this House would break them up.

1. **Elizabeth Warren** is a candidate for the President of the United States in 2020. Among other causes, she runs on a platform breaking up big tech firms such as Google and Amazon in a platform component and a supplier component. Donations go towards her campaign for the presidency — Proposition

2. **The Open Markets Institute** uses journalism to promote greater awareness of the political and economic dangers of monopolization, identifies the changes in policy and law that cleared the way for such consolidation, and fosters discussions with policymakers and citizens as to how to update America's traditional political economic principles for our 21st century digital society — Proposition

This House believes that it is in the interest of men for gender roles (e.g. work-life balance, choice of occupation, assertiveness expectations) to be dismantled

1. **The Smash Stereotypes campaign** by the Fawcett Society brings together experts across sectors to establish how society can end gender stereotyping, and to induce government action on the topic. — Proposition

2. **Jordan Peterson** is an influential psychologist who maintains that excessive feminisation of society is a major cause of anxiety in men. He accepts donations for online education and other projects. — Opposition

This House believes that humanitarian organizations should primarily use apolitical aid as opposed to politicized action

1. The **Boycott, Divestment, Sanctions** movement works to support Palestinians' struggle for justice and pressure Israel to comply with international law. — Opposition

2. **The Campaign Against Arms Trade** is a UK based humanitarian organization that takes political action to end all UK government support for arms exports, especially to Saudi Arabia, Nigeria, the US, and other countries where UK arms have been used against civilians. — Opposition

This House Would allow people to sign income share agreements

1. **The Student Freedom Initiative** is a large-scale effort to ease the disproportionate loan burden on Black students in the U.S. by offering income share agreements as a cheaper, less risky alternative to student loans. — Proposition

2. **The Student Borrower Protection Center** advocates against the use of income share agreements, which they view as a predatory lending tool that may sidestep some of the stricter regulation on student debt. Opposition

This House believes that developing nations should prioritise investment in future technologies (e.g. 5G, cybersecurity, green energy) over traditional sectors (e.g. agriculture, manufacturing)

1. **Farm Africa** is a charity that aims to reduce poverty by supporting rural agriculture in Africa. — Opposition

2. **Digital Divide Data** brings tech skills to men and women in under-served communities in Asia, offering a wide range of digital content and technology services. — Proposition

This House supports engaging private military companies to combat terrorism

1. **International Corporate Accountability Roundtable** works to protect human rights in the face of corporate abuse, including by PMCs. — Opposition

2. **War on Want** actively advocates against the widespread engagement of private military companies and other comercialisations of war. — Opposition

This House regrets modern medicine's prioritisation of longevity over quality of life (e.g. extending lifespan vs. palliative care).

1. The **Palliative Care Network** aims to improve the quality of life of patients who are facing end-of-life issues, by promoting palliative care education and collaboration globally. — Proposition

2. The **Lifespan Research Institute** aims to extend lifespan by discovering anti-aging compounds that demonstrably extend human life and prevent the onset of age-related disease. — Opposition

This House proposes a world where, after COVID-19 is no longer a threat to public health and safety, remote work remains the norm.

1. **Badass Digital Nomads** prepares people for success in the remote economy by highlighting cutting edge trends in technology and remote entrepreneurship. — Proposition

2. The **Remote Work Association** collects and distributes resources for new-to-remote employers and workers around the world to facilitate telework connections. — Proposition

This House opposes political consumerism.

1. **Progressive Shopper** makes available information on companies' donations to US political parties so that shoppers can consume in accordance with their political preferences. — Opposition

2. The **Grab your Wallet Alliance** publishes a list of companies that have a connection to Donald Trump, in order to allow consumers to flex their economic power. — Opposition

This House supports the rise of private cities in the developing world.

1. The **Startup Societies Foundation** promotes and connects startup societies (such as private cities, eco villages, and smart cities) in order to create a vibrant, global industry in this field.	Proposition
2. The **Charter Cities Institute** is a non-profit organization that supports the creation of novel governance systems by coordinating various stakeholders in new charter or private cities.	Proposition

References

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth, "Measuring and bounding experimenter demand," *American Economic Review*, 2018, *108* (11), 3266–3302.

Harrison, Glenn W, Jimmy Martínez-Correa, and J Todd Swarthout, "Eliciting subjective probabilities with binary lotteries," *Journal of Economic Behavior & Organization*, 2014, *101*, 128–140.

Imai, Kosuke, Luke Keele, and Dustin Tingley, "A general approach to causal mediation analysis.," *Psychological methods*, 2010, *15* (4), 309.

_ , _ , and Teppei Yamamoto, "Identification, inference and sensitivity analysis for causal mediation effects," *Statistical science*, 2010, pp. 51–71.