Research Article

Psychological Science
2017, Vol. 28(12) 1745–1762

DOI: 10.1177/0956797617713798
www.psychologicalscience.org/PS
SAGE

The Humanizing Voice: Speech Reveals, and Text Conceals, a More Thoughtful Mind in the Midst of Disagreement

Juliana Schroeder[1], Michael Kardas[2], and Nicholas Epley[2]
[1]Haas School of Business, University of California, Berkeley, and [2]Booth School of Business, The University of Chicago

Abstract

A person's speech communicates his or her thoughts and feelings. We predicted that beyond conveying the contents of a person's mind, a person's speech also conveys mental capacity, such that hearing a person explain his or her beliefs makes the person seem more mentally capable—and therefore seem to possess more uniquely human mental traits—than reading the same content. We expected this effect to emerge when people are perceived as relatively mindless, such as when they disagree with the evaluator's own beliefs. Three experiments involving polarizing attitudinal issues and political opinions supported these hypotheses. A fourth experiment identified paralinguistic cues in the human voice that convey basic mental capacities. These results suggest that the medium through which people communicate may systematically influence the impressions they form of each other. The tendency to denigrate the minds of the opposition may be tempered by giving them, quite literally, a voice.

Keywords

dehumanization, conflict, communication, mind perception, social cognition, open data, open materials, preregistered

Received 12/15/16; Revision accepted 5/16/17

The most basic divide in social life is between the self and others. The self is experienced from an inside perspective as a collection of ongoing mental states, including thinking, reasoning, feeling, and wanting. Others, in contrast, are experienced from an outside perspective as a collection of observed actions from which the presence of a mind is indirectly inferred (Epley & Waytz, 2010; Jones & Nisbett, 1972; Malle, Knobe, & Nelson, 2007; Pronin, 2009). "*I* think" is a fact; "*you* think" is a guess (Wegner & Gilbert, 2000).

This inferential guesswork about the minds of others is essential to social life because failing to infer that another person has mental capacities similar to one's own is the essence of dehumanization—that is, representing others as having a diminished capacity to either think or feel, as being more like an animal or an object than like a fully developed human being (Gray, Gray, & Wegner, 2007; Harris & Fiske, 2009; Haslam, 2006; Haslam, Loughnan, & Holland, 2013; Leyens et al., 2000; Waytz, Schroeder, & Epley, 2014). Such dehumanization is especially common when people evaluate an out-group member who holds beliefs, values, or attitudes different from their own (Haslam & Loughnan, 2014). Instead of attributing disagreement to different ways of thinking about the same problem, people may attribute disagreement to the other person's inability to think reasonably about the problem (Kennedy & Pronin, 2008; Pronin, Lin, & Ross, 2002). As George Carlin once wisely joked, "Have you ever noticed when you're driving that anyone who's driving slower than you is an idiot and anyone driving faster than you is a maniac?" (Carlin, 1984).

If other people's minds must be inferred, then cues connected to ongoing mental experience may be used to infer the presence of humanlike mental capacities in others. Here, we suggest that a person's voice, through speech, provides cues to the presence of thinking and

Corresponding Author:
Juliana Schroeder, University of California, Berkeley, Haas School of Business, 2220 Piedmont Ave., Berkeley, CA 94720
E-mail: jschroeder@haas.berkeley.edu

feeling, such that hearing what a person has to say will make him or her appear more humanlike than reading what that person has to say.

We base our prediction on existing theory and empirical results. Theoretically, the human voice is a tool for communicating the content of one's mind to others (Pinker & Bloom, 1990). Even when speech lacks meaningful semantic content, paralinguistic cues can convey the valence of emotional experience or intention (McAleer, Todorov, & Belin, 2014; Scherer, Banse, & Wallbott, 2001; Weisbuch, Pauker, & Ambady, 2009). A person's mental states can therefore be inferred more accurately via speech than via text (Hall & Schmid Mast, 2007; Kruger, Epley, Parker, & Ng, 2005). Beyond revealing underlying mental states, paralinguistic cues also appear to communicate humanlike mental capacities related to thinking and feeling, such as the capacity for reasoning, intellect, and emotional experience (Schroeder & Epley, 2015, 2016). Indeed, a person's voice is a social cue that may be uniquely capable of revealing his or her mental experiences related to thinking and feeling while he or she is in the midst of having those experiences. A rising pitch may convey enthusiasm. A slowed pace or pause may convey analytical reasoning. Just as variance in bodily movement (i.e., biological motion) serves as a cue for the presence of biological life, so too may variance in paralinguistic cues, such as intonation and pace, serve as a cue for the presence of an active mental life. Text alone lacks these paralinguistic cues that reveal uniquely human mental capacities, thereby enabling dehumanization if readers do not compensate for the absence of these cues.

Several empirical results suggest our hypothesis. In one series of experiments, job candidates delivering "elevator pitches" were judged to be more intelligent, thoughtful, and rational—traits consistent with perceived humanity—when evaluators heard the pitches than when they read transcripts of the same pitches or read the candidates' written pitches (Schroeder & Epley, 2015). Being able to see the candidates deliver the pitches, which provided visual cues, did not increase evaluations of the candidates' intellect. This suggests that mental capacities related to perceived humanity may be uniquely conveyed through a person's voice. In another series of experiments, participants were more likely to infer that a speech was created by a mindful human than by a mindless machine when they heard the speech being read by an actor than when they read the same semantic content, regardless of whether the speech was actually created by a human or by a computer (Schroeder & Epley, 2016). Although these experiments did not measure humanization directly, their results suggest that cues related to humanization may be conveyed through voice.

The research we report here has the potential to advance existing knowledge in four ways. First, we examined a new domain (political and social conflict) in which dehumanization is both common and consequential. Second, we advanced the developing literature on dehumanization by identifying voice as a potential moderator of dehumanization using a previously validated scale (Haslam, Bain, Douge, Lee, & Bastian, 2005). Humanization is empirically distinct from general positivity because traits perceived to distinguish humans from nonhumans can also be undesirable (e.g., impatience, jealousy; Haslam & Bain, 2007; Haslam et al., 2005). Humanization is instead a more precise form of social cognition reflecting evaluations that distinguish a person from animals (i.e., traits of *human uniqueness*, which are related to the capacity for thought and include, for example, rationality and intellect) or objects (i.e., traits of *human nature*, which are related to the capacity for emotional experience and include, for example, responsiveness and warmth). Third, we examined whether an observer's agreement with another person moderates the effect of that person's communication medium on his or her dehumanization. Our primary prediction was that hearing a person's voice would increase evaluators' attribution of human traits to that person in cases of disagreement, because this is when others are most likely to be dehumanized (perceived as irrational, illogical, or unsophisticated). In contrast, people tend to evaluate similar others by relying on egocentric projection rather than behavioral cues (Ames, 2004; Krueger, 2000), which suggests that communication medium may not reliably influence evaluations in cases of agreement. We therefore analyzed evaluators' impressions in cases of agreement and disagreement separately. Finally, we identified which paralinguistic cues humanize a speaker by comparing evaluations of human voices and computer-generated voices (Experiment 4). For all four experiments, we report how we determined our sample size, all data exclusions, all manipulations, and all measures. Attrition analyses are presented in the Supplemental Material available online.

Experiment 1: Polarizing Issues

In this experiment, we first videotaped people (*communicators*) explaining their attitude on a polarizing issue. We then asked other people (*evaluators*) to watch, listen to, or read transcripts of these explanations and to rate the communicators on several traits. Some of the evaluators agreed and others disagreed with their assigned communicators. We predicted that the media containing voice (i.e., audiovisual and audio files) would reduce the tendency to dehumanize a person with an opposing viewpoint compared with the medium lacking voice (i.e., transcript). By comparing the audiovisual and audio conditions, we further tested whether individuating cues have an additive effect on humanization, such that vocal plus visual cues are more humanizing than vocal cues alone. If the combination of visual cues and vocal cues, compared with vocal cues alone, does not increase humanization, this would suggest that humanization may be uniquely conveyed via voice. Measuring agreement allowed us to test whether or not this factor moderates the effect of communication medium on humanization.

Method

Participants

Communicators. We recruited communicators from an e-mail list of a research laboratory in downtown Chicago. Respondents completed an online pretest; we told them that they would receive $2.00 if they were selected to participate in a subsequent experiment. We precommitted to running the pretest survey for one weekend. In total, 31 people (mean age = 33.23 years, *SD* = 13.81; 61% female, 39% male) completed the survey, which asked them to report the valence and strength of their opinions on a series of potentially polarizing topics.

We then selected the three issues that yielded the most polarized responses (i.e., the largest standard deviations): abortion, the U.S. war in Afghanistan, and music (preference for country vs. rap music). Specifically, the questions regarding these issues were as follows:

- "(1) Abortion is the termination of a pregnancy by the removal or expulsion of a fetus or embryo from the uterus, resulting in or caused by its death. Which of the following options best fits your viewpoint on abortion?" (0 = *I completely oppose abortion*, 6 = *I completely support abortion*)
- "(2) The United States has been at war with Afghanistan since 2001. Which of the following options best fits your viewpoint on this war?" (0 = *I completely oppose the war*, 6 = *I completely support the war*)
- "(3) Please rate how much you enjoy country music" (0 = *do not at all enjoy*, 6 = *strongly enjoy*)
- "(4) Please rate how much you enjoy rap music" (0 = *do not at all enjoy*, 6 = *strongly enjoy*).

Finally, we selected 6 communicators (mean age = 38.3 years, *SD* = 11.2; 50% female, 50% male): the respondent on each side of each issue who had the most extreme, and strongest, opinions. Therefore, our final sample of communicators contained one person who opposed abortion, one who supported abortion, one who opposed the war, one who supported the war, one who enjoyed country music, and one who enjoyed rap music.

Evaluators. We targeted a sample of 360 evaluators in an attempt to obtain ratings from 10 evaluators for each communicator in each experimental condition. In total, we collected data from 320 Amazon Mechanical Turk workers (mean age = 32.61 years, *SD* = 11.81; 51% female, 49% male; all U.S. citizens), who participated in exchange for $1.00 each. We excluded 23 evaluators whose speed indicated that they did not pay sufficient attention to the survey (see the Results section for more details), so our final sample consisted of 297 evaluators.

Procedure

Communicators. When the 6 selected communicators returned to the laboratory, we first reminded them of their stated opinions in the pretest survey and then provided the following instructions:

> You have been selected to be in this study because of your opinions about this topic from the pre-survey. Please think carefully about your opinion and the reasons why you hold it. For the next 3 minutes, you will explain your views on this topic. Someone else will watch this video, and you should imagine you are talking directly to that person. You are trying to explain your point of view to the person, and trying to get the person to understand you. Please discuss your opinion in depth, and make sure to talk about the relevant aspects of it.

Communicators sat in a chair facing a video camera and spoke about their opinions until their speeches reached their natural conclusions (speech durations ranged from 1 to 3 min). One research assistant transcribed the speeches, and a second checked the transcriptions for accuracy. We removed verbal filler words (e.g., "um") unless their exclusion changed a sentence's meaning (in accord with the transcription method used in Schroeder & Epley, 2015, 2016).

Evaluators. Evaluators first reported their opinions on the three selected topics from the communicators' pretest so we could assess their agreement with the communicators. We then randomly assigned each evaluators to 1 of 18 conditions in a 3 (communication medium: audiovisual, audio, transcript) × 6 (communicator) between-participants design. Because the evaluators either disagreed or agreed with the communicators' opinions, this yielded a total of 36 experimental conditions.

Each evaluator then watched (audiovisual condition), listened to (audio condition), or read (transcript condition) a single speech from a communicator who either supported or opposed one of the three speech topics. Just before the stimulus was presented, participants read the following (manipulation of communication medium indicated by slashes):

> You will watch a video of/listen to/read a transcript of another participant talking about some of their opinions. Please consider their opinions as you watch/listen/read. You will be asked a few questions afterwards about the speaker and their opinions.

Throughout the survey, we referred to the communicator as the "speaker" because it was clear that the communicator was talking, rather than writing, about his or her opinions.

To measure evaluations of the communicators' humanlike capacities as comprehensively as possible, we then asked the evaluators to complete the most widely used and well-validated measure of humanization (Bastian & Haslam, 2010; Haslam & Bain, 2007; Haslam et al., 2005), plus additional items measuring perceived mental capacities of thinking and feeling (e.g., perceived thoughtfulness, rationality, emotionality, and likeability; see the Supplemental Material). Because the results for the additional measures were consistent with those for the humanization scale, and to keep our discussion focused on our primary prediction, we present only the results for the latter scale here. Results for the additional items are presented in the Supplemental Material.

We used the humanization scale developed by Bastian and Haslam (2010), which measures two dimensions of humanization. The Human Uniqueness subscale includes 6 items generally related to higher-order cognition and intellectual competence: Evaluators rated the extent to which the speaker was "refined and cultured"; was "rational and logical"; lacked "self-restraint" (reverse-scored); was "unsophisticated" (reverse-scored); was "like an adult, not a child"; and seemed "less than human, like an animal" (reverse-scored). The Human Nature subscale includes 6 items generally related to emotional experience and interpersonal warmth: Evaluators rated the extent to which the speaker was "open-minded"; was "emotional, responsive, and warm"; was "superficial" and lacked "depth" (reverse-scored); was "mechanical and cold, like a robot" (reverse-scored); was "like an object, not a human" (reverse-scored); and had "interpersonal warmth." All 12 items were presented with response scales ranging from −3 (*much less than the average person*) to 3 (*much more than the average person*). Cronbach's α was .82 for the Human Uniqueness subscale and .83 for the Human Nature subscale.

Finally, evaluators completed a memory test intended to capture any possible differences in attention to the communicator's speech across conditions. If we found greater humanization of the communicators in the audiovisual and audio conditions than in the transcript condition, this could be interpreted as indicating that the speeches in the audiovisual and audio conditions were more engaging, and therefore more memorable, than the speeches in the transcript condition. To assess this possibility, we asked the evaluators to "please write as much as you can remember about the speaker and his or her experience" in a text box. Because we did not find evidence consistent with this alternative interpretation, and did not obtain consistent results for it across our four experiments, we do not discuss the memory test further here, but the results for this test are included in the Supplemental Material.

Results

The final sample consisted of 297 evaluators after we excluded 23 evaluators whose speed indicated that they could not possibly have watched, listened to, or read the explanations fully (less than 60 s in the audiovisual and audio conditions and less than 20 s in the transcript condition). Eight evaluators were excluded from the audiovisual condition, 4 from the audio condition, and 11 from the transcript condition. Exclusions did not vary by experimental condition, $\chi^2(2, N = 320) = 3.07$, $p = .215$. We observed no statistically significant interactions between communicator's topic and communication-medium condition for any of our dependent variables (including those discussed in the Supplemental Material), Fs < 0.25, ps $> .250$, and therefore did not include communicator's topic as a variable in the analyses reported here.

To distinguish evaluations in cases of disagreement from evaluations in cases of agreement, we coded the evaluators according to their self-reported opinion on the topic to which they were assigned. Those with scores below 3 (on the scale from 0 to 6) were coded as disagreeing with a communicator who spoke in favor of the topic and as agreeing with a communicator who spoke against the topic. In contrast, evaluators with scores above 3 were coded as disagreeing with a communicator who spoke against the topic and as agreeing with a communicator who spoke in favor of the topic. Evaluators who rated their opinions exactly at the midpoint of the scale (3) were always coded as disagreeing with the communicator, to be consistent with our coding in other experiments. However, coding these evaluators ($n = 51$) at the midpoint as agreeing with the communicator did not meaningfully alter the results of any of the analyses we report here (see the Supplemental Material). We excluded from all analyses 1 evaluator who did not report his or her opinion on one item and whose agreement could therefore not be coded (thus, 296 participants were included in analyses). In total, 178 participants disagreed with (or were neutral toward) their assigned communicator's opinions, and 118 participants agreed. Agreement did not

Table 1. Descriptive Statistics for the Primary Comparisons in the Four Experiments

Experiment and communication medium	Evaluators who disagreed with the communicator						Evaluators who agreed with the communicator					
	Ratings of human uniqueness			Ratings of human nature			Ratings of human uniqueness			Ratings of human nature		
	M	*SD*	95% CI	*M*	*SD*	95% CI	*M*	*SD*	95% CI	*M*	*SD*	95% CI
Experiment 1												
Audiovisual	0.82	1.08	[0.56, 1.09]	0.82	1.04	[0.56, 1.08]	0.76	0.89	[0.46, 1.07]	0.84	0.94	[0.51, 1.17]
Audio	0.63	0.97	[0.38, 0.88]	0.71	1.08	[0.43, 0.99]	1.01	1.02	[0.67, 1.34]	0.96	0.93	[0.66, 1.27]
Transcript	0.06	1.11	[−0.25, 0.37]	0.27	1.18	[−0.06, 0.60]	0.73	1.12	[0.39, 1.06]	1.15	1.04	[0.84, 1.46]
Experiment 2												
Audiovisual	0.80	0.94	[0.61, 0.99]	0.70	1.03	[0.49, 0.91]	1.22	1.04	[0.92, 1.53]	1.08	0.92	[0.81, 1.35]
Audio	0.63	1.19	[0.39, 0.86]	0.48	1.01	[0.28, 0.68]	1.15	0.99	[0.88, 1.43]	1.21	0.97	[0.94, 1.47]
Transcript	0.09	1.26	[−0.16, 0.34]	0.36	0.97	[0.17, 0.56]	0.94	1.04	[0.66, 1.22]	1.16	1.07	[0.87, 1.45]
Written	0.30	1.09	[0.08, 0.52]	0.30	0.97	[0.10, 0.49]	0.82	1.08	[0.53, 1.11]	1.01	0.98	[0.74, 1.27]
Experiment 3												
Audiovisual	0.70	1.28	[0.48, 0.91]	0.52	1.21	[0.31, 0.72]	1.44	1.06	[1.19, 1.69]	1.29	1.17	[1.01, 1.57]
Audio	0.45	1.24	[0.25, 0.65]	0.26	1.26	[0.06, 0.47]	1.59	0.87	[1.39, 1.79]	1.36	0.85	[1.16, 1.56]
Transcript	0.07	1.23	[−0.15, 0.28]	0.13	1.19	[−0.08, 0.33]	0.91	1.02	[0.68, 1.14]	1.04	1.00	[0.82, 1.27]
Written	0.15	1.30	[−0.06, 0.36]	−0.05	1.17	[−0.24, 0.14]	1.29	0.95	[1.06, 1.53]	1.06	1.06	[0.79, 1.32]
Experiment 4												
Authentic voice	0.66	1.11	[0.48, 0.84]	0.37	1.00	[0.20, 0.53]	0.99	1.11	[0.70, 1.28]	0.69	1.00	[0.43, 0.96]
Mindless voice	0.49	1.09	[0.31, 0.66]	−0.04	1.19	[−0.23, 0.15]	1.16	1.13	[0.85, 1.46]	0.59	1.43	[0.21, 0.98]
Transcript	0.28	1.14	[0.09, 0.48]	0.41	0.94	[0.25, 0.57]	0.82	1.18	[0.58, 1.07]	1.07	1.16	[0.83, 1.32]

Note: CI = confidence interval.

vary by communication medium, $\chi^2(2, N = 296) = 3.53$, $p = .171$, or by communicator's topic, $\chi^2(2, N = 296) = 0.50$, $p > .250$.

Table 1 summarizes the ratings of human uniqueness and human nature in the three communication-medium conditions, separately for evaluators who agreed and those who disagreed with their assigned communicators. We conducted planned contrasts to test our primary prediction that evaluators would dehumanize communicators with an opposing viewpoint less when they heard what the communicator had to say than when they read it. As we predicted, among evaluators who disagreed with their communicators, communication medium significantly affected ratings of both communicators' human uniqueness, $F(2, 175) = 8.06$, $p < .001$, $\eta^2 = .08$, and their human nature, $F(2, 175) = 3.91$, $p = .022$, $\eta^2 = .04$ (see Fig. 1). Specifically, evaluators in the audio condition judged communicators who disagreed with them to be significantly more humanlike (human uniqueness: $M = 0.63$, $SD = 0.97$, 95% confidence interval, CI = [0.38, 0.88]; human nature: $M = 0.71$, $SD = 1.08$, 95% CI = [0.43, 0.99]) than did evaluators in the transcript condition (human uniqueness: $M = 0.06$, $SD = 1.11$, 95% CI = [−0.25, 0.37]; human nature: $M = 0.27$, $SD = 1.18$, 95% CI = [−0.06, 0.60]), and this effect was found for evaluations of both human uniqueness, $t(175) = 2.88$, $p = .004$, $d = 0.55$, and human nature, $t(175) = 2.12$, $p = .036$, $d = 0.40$. Evaluators in

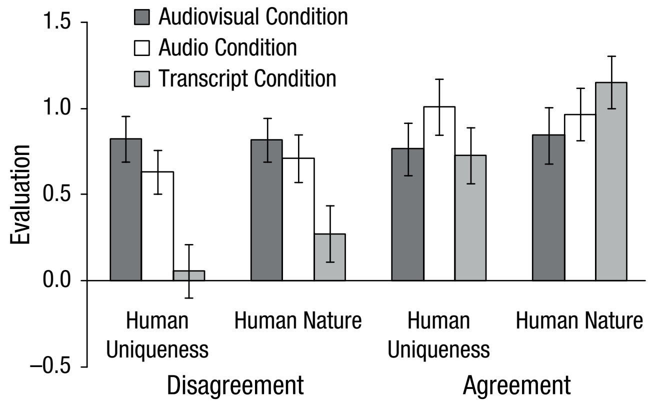

Fig. 1. Evaluations of communicators' human-uniqueness and human-nature traits in the audiovisual, audio, and transcript conditions of Experiment 1. Results are presented separately for evaluators who agreed and who disagreed with the communicators they rated. Error bars represent ±1 *SEM*.

the audiovisual condition also judged communicators to be significantly more humanlike (human uniqueness: $M = 0.82$, $SD = 1.08$, 95% CI = [0.56, 1.09]; human nature: $M = 0.82$, $SD = 1.04$, 95% CI = [0.56, 1.08]) than did evaluators in the transcript condition, and again, this effect was found for evaluations of both human uniqueness, $t(175) = 3.92$, $p < .001$, $d = 0.73$, and human nature, $t(175) = 2.69$, $p = .008$, $d = 0.50$. We observed no significant difference between the audiovisual and audio condition in evaluations of communicators' human uniqueness, $t(175) = 1.01$, $p > .250$, $d = 0.18$, or human nature, $t(175) = 0.55$, $p > .250$, $d = 0.10$. Thus, the addition of visual information did not meaningfully affect the degree to which people humanized someone with a different opinion. In contrast, when the evaluators agreed with the communicators, we observed no significant effect of communication medium on ratings of communicators' human uniqueness (see Fig. 1), $F(2, 115) = 0.87$, $p > .250$, $\eta^2 = .02$, or human nature, $F(2, 115) = 1.01$, $p > .250$, $\eta^2 = .02$.

To examine this overall pattern, we conducted a 3 (communication medium: audiovisual, audio, or transcript) × 2 (agreement: evaluator agreed or disagreed with the communicator) × 2 (measure: human-uniqueness or human-nature traits) mixed-model analysis of variance (ANOVA). This analysis revealed effects of medium, $F(2, 290) = 2.37$, $p = .095$, $\eta_p^2 = .02$; agreement, $F(1, 290) = 9.23$, $p = .003$, $\eta_p^2 = .03$; and measure, $F(1, 290) = 8.59$, $p = .004$, $\eta_p^2 = .03$. These main effects were qualified by an interaction between agreement and medium, $F(2, 290) = 3.82$, $p = .023$, $\eta_p^2 = .03$, and an interaction between measure and medium, $F(2, 290) = 5.39$, $p = .005$, $\eta_p^2 = .04$. All other interactions were nonsignificant, Fs < 1.33, ps > .250, η_p^2s < .01. The interaction between agreement and communication medium indicated that communication medium influenced evaluations more in cases of disagreement than in cases of agreement, as already discussed. This result led us to predict that agreement might moderate the effect of communication medium on dehumanization in the subsequent experiments as well. As we describe later, this predicted moderation was not consistently supported. The interaction between measure and communication medium indicated that communication medium influenced evaluations of human uniqueness more than evaluations of human nature. We did not anticipate this interaction, although it emerged in Experiments 2 through 4 as well. It suggests that humanlike capacities related to thinking and cognition (those measured by human uniqueness) may be conveyed more clearly over voice than capacities related to emotional experience and interpersonal warmth (those measured by human nature). We address this interesting possibility in the General Discussion.

Experiment 2: Polarizing Political Primaries

Experiment 1 suggested that when people evaluate a person with an opposing viewpoint, they may humanize that person more if they hear the person's voice than if they read what he or she has written. Experiment 2 tested this hypothesis in another context in which people routinely derogate those with opposing views: political elections. People recruited during the 2016 U.S. presidential primaries explained why they preferred their chosen candidate in speech and writing. By including communicators' own written explanations (which they typed), we tested whether the results of Experiment 1 were due to reading speech transcriptions rather than to the absence of human voice in text. We predicted that evaluators who heard a voter's opposing viewpoint would humanize the voter more than would those who read the voter's opposing viewpoint.

Method

Participants

Communicators. As in Experiment 1, we included multiple communicators to increase the generalizability and ecological validity of the experiment (Wells & Windschitl, 1999). Our goal was to recruit equal numbers of voters supporting Democratic and Republican candidates. Using an online announcement posted to the same e-mail list as in Experiment 1, we recruited 4 Democratic and 4 Republican communicators (mean age = 35.38 years, $SD = 12.68$; 25% female, 75% male) who were willing to discuss their preferred candidate in exchange for $10.

Evaluators. We targeted a sample of 640 evaluators in an attempt to obtain at least 10 evaluations for each communicator in each experimental condition. A total of 643 Mechanical Turk workers (mean age = 35.15 years, SD = 11.73; 49% female, 51% male; all U.S. citizens) completed the online survey in exchange for $0.75 each.

Procedure

Communicators. The 4 Democratic and 4 Republican communicators visited the laboratory and first responded to two questions: "Which candidate do you support for the 2016 U.S. Presidential election?" (free response) and "What is this candidate's political party?" ("Democratic," "Republican," "other"). We conducted this experiment during the U.S. presidential primaries, when there were still multiple candidates competing for their party's nomination. The communicators supported the following candidates: Bernie Sanders (Democrat; n = 3), Hillary Clinton (Democrat; n = 1), John Kasich (Republican; n = 3), and Donald Trump (Republican; n = 1).

Each communicator both spoke and wrote about the reasons for his or her support (order counterbalanced). The experimenter provided the following instructions (manipulation of communication medium indicated by slashes):

> We are recruiting participants from various political backgrounds because we are interested in understanding people's political beliefs and how people communicate those beliefs to others. Now that you have reported which candidate you support for the upcoming U.S. Presidential election, we would like you to speak/write about why you support this candidate. [*Speaking condition only:* You will speak out loud while we video record your response.] We will show the recording/what you write to another study participant who may have similar or different political beliefs and we would like you to imagine that you are speaking/writing directly to the study participant. Please think carefully about your opinion and the reasons why you hold it. Try to explain your point of view to the study participant, and try to get that person to understand you. Please discuss your opinions in depth, such as why you support this candidate, why you prefer this candidate over other candidates and what you like about this candidate. First, please jot down notes on this sheet and then tell me when you are ready to begin recording.

After receiving these instructions, the communicators both spoke about the reasons for their support, while seated in a chair facing a video camera, and wrote about the reasons for their support, while seated in front of a laptop. We allowed them to speak and write for as long as they wanted. They spent between 40 s and 3 min speaking, and between 2 min and 12 min writing. We observed a statistically nonsignificant difference in the number of words spoken versus written (spoken: M = 251.75, SD = 112.32, 95% CI = [157.85, 345.65]; written: M = 202.50, SD = 123.34, 95% CI = [99.38, 305.62]), paired-samples $t(7)$ = 1.84, p = .108, d = 0.65.

One research assistant transcribed the speeches, and a second checked for accuracy. As in Experiment 1, we removed verbal filler words from the transcripts (e.g., "uh"), unless their exclusion changed the sentence's meaning. For the written condition, we did not make any changes to the communicators' written texts, just as we did not make any changes to the spoken stimuli in the audio and audiovisual conditions.

Evaluators. In this experiment, we included three attention checks designed to identify evaluators who were not paying adequate attention to the stimuli so that they could be excluded from all analyses. The first attention check came at the beginning of the survey. All evaluators watched a short audiovisual test clip and reported what they saw and heard so that we could exclude evaluators who misreported the video's content. The other two attention checks came at the end of the experiment. All evaluators were asked, "Did you pay attention throughout the whole study?" ("yes" or "no") and "In what form did we show you the participant's opinions?" ("video," "audio," "transcript," "written"). Evaluators who answered "no" to the first question or who mistook a voice condition (i.e., video or audio) for a text condition (i.e., transcript or written), or vice versa, were excluded from analysis.

After completing the first attention check, the evaluators reported their political-party affiliation (0 = *I completely support the Democratic party*, 3 = *not sure/I am politically moderate*, 6 = *I completely support the Republican party*) and how strongly they felt about the topic (0 = *I don't care at all*, 3 = *not sure*, 6 = *I feel extremely strongly*). Then they reported which candidate they supported for the upcoming U.S. presidential election and how favorably they viewed each of the candidates (0 = *extremely unfavorable*, 6 = *extremely favorable*). At the time (i.e., in the midst of the states' 2016 primary elections), the Democratic and Republican presidential candidates with the highest polling numbers (based on aggregate polling data within the respective parties) were Donald Trump (Republican), John Kasich (Republican), Ted Cruz (Republican), Hillary Clinton (Democrat), and Bernie Sanders (Democrat). We presented these five candidates to the evaluators in randomized order. Because there was intense disagreement both within and between the political parties at this particular time, we measured agreement using each evaluator's favorability rating of the assigned communicator's preferred candidate rather than party affiliation.

We randomly assigned the evaluators to 32 experimental conditions in a 4 (communication medium: audiovisual, audio, transcript, written) × 8 (communicator) between-participants design. Some evaluators perceived the assigned communicator's candidate choice favorably, and others perceived the assigned communicator's candidate choice unfavorably, so there were 64 unique experimental conditions. Each evaluator watched and listened to the videotaped speech (audiovisual condition), listened to the speech only (audio condition), read the transcribed speech (transcript condition), or read the written statement (written condition) of a single communicator with whom he or she either agreed or disagreed regarding choice of the presidential candidate. To ensure that evaluators in the audio and audiovisual conditions observed the assigned communicator's entire statement, we programmed the survey so that the clips automatically paused if an evaluator clicked outside of the window containing the audio or video player.

After watching, listening to, or reading the speech or reading the written statement, the evaluators completed the same humanization scale as in Experiment 1 (α = .88 for human uniqueness and .86 for human nature), along with a battery of other items (similar to those used in Experiment 1) measuring inferences about the communicators' mental capacities (see the Supplemental Material). The evaluators then completed four exploratory items designed to measure the communicator's persuasiveness: (a) "How much do you think your beliefs have changed as a result of the participant?" (0 = *no change*, 6 = *a lot of change*); (b) "How persuasive did you find the participant's message?" (0 = *not at all*, 6 = *very*); (c) "How hard did the participant think about their beliefs?" (0 = *not at all*, 6 = *extremely*); and (d) "How rational are the participant's beliefs?" (0 = *not at all*, 6 = *very*). We suspected that humanizing someone with an opposing viewpoint might lead evaluators to find his or her beliefs to be more reasonable, and therefore that evaluators would feel more persuaded by the communicator's explanation if they heard rather than read it. Finally, the evaluators reported their demographic information.

Results

Our final sample consisted of 607 evaluators, after exclusion of the 36 evaluators who failed one or more of the attention checks. We excluded 10 evaluators in the audiovisual condition, 5 in the audio condition, 8 in the transcript condition, and 13 in the written condition. The number of exclusions did not differ by communication medium, $\chi^2(3, N = 643) = 3.37$, $p = .338$.

To test our primary hypothesis, we first coded evaluators who rated their opinion of the assigned communicator's selected presidential candidate as 3 or less as disagreeing with the communicator (n = 395) and those who rated their opinion of that candidate as 4 or more as agreeing with the communicator (n = 212). This coding system was consistent with the system used in Experiment 1. Agreement with the communicator did not vary by communication medium, $\chi^2(3, N = 607) = 0.58$, $p = .902$.

Table 1 summarizes the ratings of human uniqueness and human nature in the four communication-medium conditions, separately for evaluators who agreed and who disagreed with their assigned communicators. As in Experiment 1, evaluators dehumanized a communicator with an opposing viewpoint less when they heard what that communicator had to say than when they read it. Communication medium again affected evaluations of both communicators' human uniqueness, $F(3, 391) = 7.93$, $p < .001$, $\eta^2 = .06$, and communicators' human nature, $F(3, 391) = 3.03$, $p = .029$, $\eta^2 = .02$. Evaluations of communicators' human uniqueness did not differ significantly between the two conditions with voice—the audiovisual condition (M = 0.80, SD = 0.94, 95% CI = [0.61, 0.99]) and the audio condition (M = 0.63, SD = 1.19, 95% CI = [0.39, 0.86]), $t(391) = 1.10$, $p = .274$, $d = 0.16$—or between the two conditions with text—the transcript condition (M = 0.09, SD = 1.26, 95% CI = [−0.16, 0.34]) and the written condition (M = 0.30, SD = 1.09, 95% CI = [0.08, 0.52]), $t(391) = -1.30$, $p = .193$, $d = -0.19$. Evaluations of communicators' human nature likewise did not differ significantly between the two conditions with voice—the audiovisual condition (M = 0.70, SD = 1.03, 95% CI = [0.49, 0.91]) and the audio condition (M = 0.48, SD = 1.01, 95% CI = [0.28, 0.68]), $t(391) = 1.54$, $p = .124$, $d = 0.22$—or between the two conditions with text—the transcript condition (M = 0.36, SD = 0.97, 95% CI = [0.17, 0.56]) and the written-statement condition (M = 0.30, SD = 0.97, 95% CI = [0.10, 0.49]), $t(391) = 0.47$, $p > .250$, $d = 0.07$. We therefore combined the data across the two conditions with voice and across the two conditions with text to test our specific hypothesis that voice diminishes the tendency to dehumanize a person with an opposing viewpoint.

As predicted, evaluators in the voice conditions judged a communicator who disagreed with them to be significantly more humanlike (human uniqueness: M = 0.71, SD = 1.07, 95% CI = [0.56, 0.86]; human nature: M = 0.59, SD = 1.03, 95% CI = [0.44, 0.73]) than did evaluators in

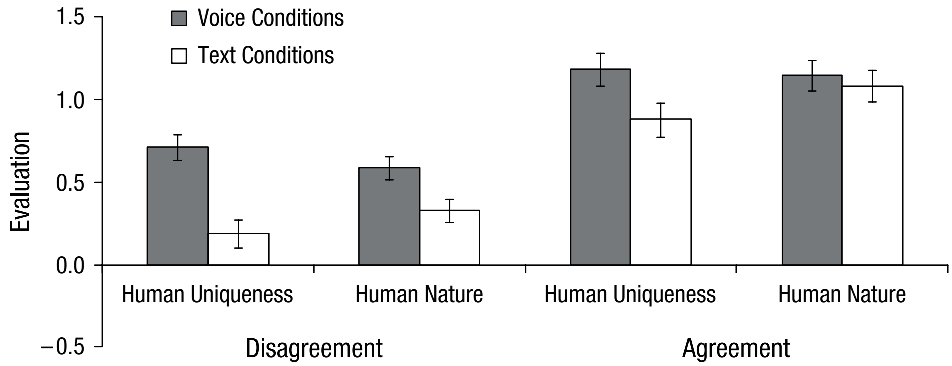

Fig. 2. Evaluations of communicators' human-uniqueness and human-nature traits in the voice conditions (audio and audiovisual) and the text conditions (transcript and written) of Experiment 2. Results are presented separately for evaluators who agreed and who disagreed with the communicators they rated. Error bars represent ±1 *SEM*.

the text conditions (human uniqueness: $M = 0.19$, $SD = 1.18$, 95% CI = [0.03, 0.36]; human nature: $M = 0.33$, $SD = 0.97$, 95% CI = [0.19, 0.47]), $t(393) = 4.57$, $p < .001$, $d = 0.46$, for human uniqueness and $t(393) = 2.55$, $p = .011$, $d = 0.26$, for human nature (see Fig. 2). In cases of agreement, we observed an unexpected difference between the voice and text conditions for ratings of human uniqueness, $t(210) = 2.15$, $p = .033$, $d = 0.30$, but not ratings of human nature, $t(210) = 0.46$, $p > .250$, $d = 0.06$ (see Fig. 2).

To test this overall pattern, we conducted a 2 (communication medium: text or voice) × 2 (agreement: evaluator agreed or disagreed with the communicator) × 2 (measure: human-uniqueness or human-nature traits) mixed-model ANOVA. This analysis revealed significant main effects of communication medium, $F(1, 603) = 11.56$, $p < .001$, $\eta_p^2 = .02$, and agreement, $F(1, 603) = 54.17$, $p < .001$, $\eta_p^2 = .08$. The effect of medium was qualified by a significant measure-by-medium interaction, $F(1, 603) = 18.01$, $p < .001$, $\eta_p^2 = .03$. The interaction between communication medium and agreement was nonsignificant, $F(1, 603) = 1.45$, $p = .228$, $\eta_p^2 < .01$. All other interactions and the remaining main effect were also nonsignificant, *F*s < 2.24, *p*s > .135, η_p^2s < .01. The interaction between measure and communication medium again indicates that the medium of communication influenced evaluations of human uniqueness more than evaluations of human nature. The nonsignificant interaction between communication medium and agreement indicates that voice was somewhat humanizing in this experiment (at least in perceptions of human uniqueness) even in cases of agreement. Although the moderating effects of agreement were in the same direction as we observed in Experiment 1, these results suggest that agreement may not be a robust moderator. Experiments 3 and 4 provided further tests of this potential moderator.

To examine whether communication medium also influenced persuasion during instances of disagreement, we combined the four persuasion items into a single index ($\alpha = .80$). Persuasion ratings did not differ significantly between the two conditions with voice—the audiovisual condition ($M = 2.77$, $SD = 1.23$, 95% CI = [2.57, 2.98]) and the audio condition ($M = 2.73$, $SD = 1.39$, 95% CI = [2.51, 2.95]), $t(603) = 0.28$, $p > .250$—or between the two conditions with text—the transcript condition ($M = 2.39$, $SD = 1.32$, 95% CI = [2.18, 2.60]) and the written-statement condition ($M = 2.44$, $SD = 1.31$, 95% CI = [2.23, 2.66]), $t(603) = -0.37$, $p > .250$. We therefore combined the data across the two conditions with voice and across the two conditions with text.

Our primary interest involved cases of disagreement. Results indicated that the evaluators were more persuaded by a speaker with whom they disagreed in the voice conditions ($M = 2.49$, $SD = 1.33$, 95% CI = [2.30, 2.67]) than in the text conditions ($M = 2.10$, $SD = 1.28$, 95% CI = [1.92, 2.28]), $t(393) = 2.92$, $p = .004$, $d = 0.29$. We observed a similar effect in cases of agreement; evaluators reported being marginally more persuaded in the voice conditions ($M = 3.28$, $SD = 1.11$, 95% CI = [3.06, 3.50]) than in the text conditions ($M = 2.97$, $SD = 1.19$, 95% CI = [2.74, 3.19]), $t(210) = 1.95$, $p = .053$, $d = 0.27$. A 2 (communication medium: voice or text) × 2 (agreement: evaluator agreed or disagreed with the communicator) ANOVA on persuasion revealed an unsurprising effect of agreement, $F(1, 603) = 60.05$, $p < .001$, $\eta_p^2 = .09$; evaluators reported being more persuaded by communicators with whom they agreed ($M = 3.12$, $SD = 1.16$) than by communicators with whom they disagreed ($M = 2.30$, $SD = 1.32$). We also observed a significant effect of communication medium, $F(1, 603) = 10.49$, $p = .001$, $\eta_p^2 = .02$; evaluators in the voice conditions reported being more persuaded ($M = 2.75$, $SD = 1.31$) than evaluators in the text conditions

($M = 2.42$, $SD = 1.32$). The medium-by-agreement interaction was nonsignificant, $F(1, 603) = 0.12$, $p > .250$, $\eta_p^2 < .01$.

To better understand the relationship between persuasion and humanization, we conducted an exploratory analysis testing whether evaluations of humanlike traits mediated the effect of being in a voice-based (vs. text-based) medium on persuasion. To simplify this analysis, we created a single index of perceived humanness using all 12 items from the human-uniqueness and human-nature scales ($\alpha = .92$). A 5,000-sample bootstrapped mediation model (Preacher & Hayes, 2008) indicated that evaluations of humanlike traits fully mediated the effect of medium on persuasion in cases of disagreement, indirect effect = 0.34, $SE = 0.09$, 95% CI = [0.17, 0.52]. However, evaluations of humanlike traits did not mediate the effect of medium on persuasion in cases of agreement, bootstrapped indirect effect = 0.13, $SE = 0.09$, 95% CI = [−0.05, 0.32]. These results are intriguing but also preliminary. Further research is necessary to understand the relationship between humanization and persuasion.

Experiment 3: Polarizing Presidential Election

Presidential primaries can polarize the electorate, leading people to denigrate opponents' minds, but general elections are often even more polarizing. To test whether our results would emerge on the cusp of an especially divisive election, we conducted a simplified replication of Experiment 2 on the weekend before the 2016 U.S. presidential election.

Method

Participants

Communicators. We recruited communicators using online announcements posted to two university laboratory pools and Craigslist, aiming for equal numbers of supporters of Hillary Clinton (the Democratic nominee for U.S. president) and Donald Trump (the Republican nominee for U.S. president). In total, we recruited 5 Clinton supporters and 5 Trump supporters (mean age = 41.60 years, $SD = 12.36$; 10% female, 90% male) who agreed to discuss their preferred candidate in exchange for \$4 plus entry into a \$100 lottery.

Evaluators. We targeted a sample of 800 evaluators, in an attempt to obtain at least 10 evaluators for each communicator in each experimental condition (10 communicators were evaluated in four communication-medium conditions by evaluators who either disagreed or agreed with the assigned communicator's preferred presidential candidate). Given the number of evaluators who failed the attention checks in Experiment 2, we anticipated excluding some evaluators from this experiment as well, and consequently recruited more than our target number to ensure that we would achieve our planned sample. Our final sample consisted of 953 Mechanical Turk workers (mean age = 36.16 years, $SD = 11.96$; 51% female, 49% male; all U.S. citizens), who served as evaluators in exchange for \$0.75 each.

Procedure

Communicators. When the 5 Clinton supporters and 5 Trump supporters visited the laboratory, they confirmed that they still intended to vote for Clinton and Trump by answering the question, "Which candidate do you plan to vote for in the 2016 U.S. Presidential election?" ("Hillary Clinton," "Donald Trump," "Gary Johnson," "Jill Stein," or "none of the above").[1] Each communicator both spoke and wrote about the reasons for his or her choice (order counterbalanced). The experimenter provided the following instructions (manipulation of communication medium indicated by slashes):

> Next we would like you to explain why you support this candidate in the election. We will video record your response/You will type your response into the computer and we will show this recording/response to another study participant who may have similar or different political beliefs from your own. We would like you to imagine that you are speaking/writing directly to the study participant. Please think carefully about your opinion and the reasons why you hold it. Try to explain your point of view to the study participant, and try to get that person to understand you. Discuss your opinions in depth, such as why you support this candidate, why you prefer this candidate over other major candidates and what you like about this candidate. Please do not reveal personally identifying information or any illegal or embarrassing activity in your response. First, please jot down notes on this sheet and then tell me when you are ready to begin recording/writing.

After receiving these instructions and writing their notes, the communicators both spoke about the reasons for their support, while seated in a chair facing a video camera, and wrote about the reasons for their support, while seated in front of a laptop. We allowed them to speak and write as long as they wanted. They spent between 48 s and 4 min speaking, and between 2 min and 14 min writing. The communicators spoke more words than they wrote (spoken: M = 306.00, SD = 113.71, 95% CI = [224.66, 387.34]; written: M = 132.10, SD = 82.21, 95% CI = [73.29, 190.91]), paired-samples $t(9)$ = 4.90, p < .001, d = 1.55. Research assistants then transcribed the speeches and checked the transcriptions for accuracy. As in Experiments 1 and 2, we removed verbal filler words from the transcripts (e.g., "uh") unless they were necessary for comprehension. For the written-statement condition, we did not make any changes to the communicators' written text, just as we did not make changes to spoken stimuli in the audio and audiovisual conditions.

Evaluators. We included four attention checks designed to identify evaluators who were not paying adequate attention to the stimuli so that they could be excluded prior to analysis. The first attention check came at the beginning of the survey. All evaluators watched a short audiovisual test clip and reported what they saw and heard so that we could exclude evaluators who misreported the video's content. The other three attention checks came at the end of the survey. We explicitly asked the evaluators, "Please tell us honestly: did you pay attention throughout the whole study?" ("yes" or "no"). We then asked them, "In what form did we show you the participant's opinions?" ("video," "audio," "transcript," "written") and "Which candidate did the participant plan to vote for?" ("Hillary Clinton," "Donald Trump"). Evaluators who answered "no" to the first question, who mistook a voice condition for a text condition or vice versa, or who answered the third question incorrectly were removed from analysis.

After completing the first attention check, the evaluators reported their political-party affiliation (0 = *I completely support the Democratic party*, 3 = *not sure/I am politically moderate*, 6 = *I completely support the Republican party*) and how strongly they felt about the topic (0 = *I don't care at all*, 3 = *not sure*, 6 = *I feel extremely strongly*). Then they reported which candidate they supported in the upcoming 2016 U.S. presidential election and how favorably they viewed each of the major candidates (presented in randomized order): Hillary Clinton, Donald Trump, Gary Johnson, and Jill Stein (0 = *extremely unfavorable*, 6 = *extremely favorable*). We used how favorably each evaluator viewed the assigned communicator's preferred candidate as our measure of agreement with the communicator.

We randomly assigned the evaluators to 40 experimental conditions in a 4 (communication medium: audiovisual, audio, transcript, written) × 10 (communicator) between-participants design. Some evaluators perceived the assigned communicator's candidate choice favorably, and others perceived the assigned communicator's candidate choice unfavorably, so there were 80 unique experimental conditions. Each evaluator watched and listened to the videotaped speech (audiovisual condition), listened to the speech only (audio condition), read the transcribed speech (transcript condition), or read the written statement (written condition) of a single communicator with whom he or she either agreed or disagreed regarding choice of the presidential candidate. After watching, listening to, or reading the speech or reading the written statement, the evaluators rated the communicators on the traits of human uniqueness and human nature only, using the same scales as in Experiment 1 (α = .89 for human uniqueness and .88 for human nature). Finally, the evaluators reported their demographic information.

Results

We excluded from analysis 102 evaluators who failed one or more of the attention checks, which left a final sample of 851 evaluators. We excluded 31 evaluators in the audiovisual condition, 17 in the audio condition, 26 in the transcript condition, and 28 in the written-statement condition). The number of exclusions did not differ by communication condition, $\chi^2(3, N = 953)$ = 4.70, p = .195. Using the same coding system as in Experiments 1 and 2, we identified 565 participants as disagreeing with their assigned communicator and 286 as agreeing with their assigned communicator. Agreement did not vary by communication medium, $\chi^2(3, N = 851)$ = 2.96, p > .250.

Table 1 summarizes the ratings of human-uniqueness and human-nature traits in the four communication-medium conditions, separately for evaluators who agreed and who disagreed with their assigned communicators. As in Experiments 1 and 2, among the evaluators who disagreed with communicators, communication medium affected evaluations of both communicators' human uniqueness, $F(3, 561)$ = 7.16, p < .001, η^2 = .04, and their human nature, $F(3, 561)$ = 5.57, p = .001, η^2 = .03. Evaluations of communicators' human uniqueness did not differ significantly between the two conditions with voice—the audiovisual condition (M = 0.70, SD = 1.28, 95% CI = [0.48, 0.91]) and the audio condition (M = 0.45, SD = 1.24, 95% CI = [0.25, 0.65]), $t(561)$ = 1.66, p = .098, d = 0.20—or between the two conditions with text—the transcript condition (M = 0.07, SD = 1.23, 95% CI = [−0.15, 0.28]) and the written-statement condition (M = 0.15, SD = 1.30, 95% CI = [−0.06, 0.36]), $t(561)$ = −0.57, p = .568, d = 0.07. Evaluations of communicators' human nature likewise did not differ significantly between the two conditions with

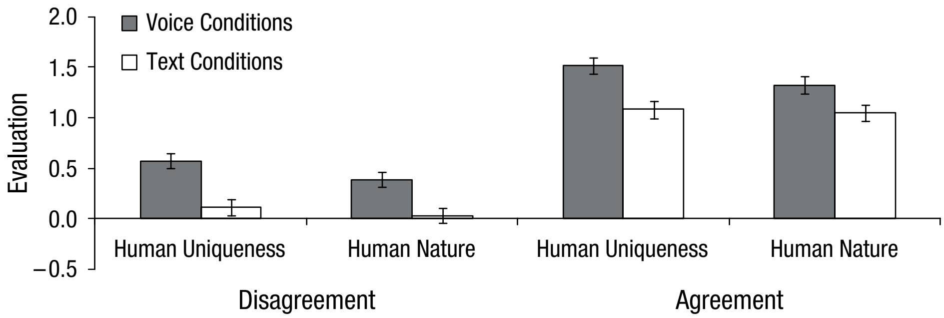

Fig. 3. Evaluations of communicators' human-uniqueness and human-nature traits in the voice conditions (audio and audiovisual) and the text conditions (transcript and written) of Experiment 3. Results are presented separately for evaluators who agreed and who disagreed with the communicators they rated. Error bars represent ±1 *SEM*.

voice—the audiovisual condition (M = 0.52, SD = 1.21, 95% CI = [0.31, 0.72]) and the audio condition (M = 0.26, SD = 1.26, 95% CI = [0.06, 0.47]), $t(561)$ = 1.77, p = .078, d = 0.12—and the two conditions with text—the transcript condition (M = 0.13, SD = 1.19, 95% CI = [−0.08, 0.33]) and the written-statement condition (M = −0.05, SD = 1.17, 95% CI = [−0.24, 0.14]), $t(561)$ = 1.19, p = .234, d = 0.14. We therefore combined the data across the two voice conditions and across the two text conditions in subsequent analyses, as we did in Experiment 2.

As predicted, evaluators in the voice conditions again judged communicators who disagreed with them to be significantly more humanlike than did evaluators in the text conditions, both in human uniqueness, $t(563)$ = 4.29, p < .001, d = 0.36, and in human nature, $t(563)$ = 3.48, p < .001, d = 0.29 (see Fig. 3). Unlike in Experiments 1 and 2, however, when evaluators agreed with their communicators, communication medium (voice conditions vs. text conditions) significantly affected evaluations of human uniqueness, $t(284)$ = 3.71, p < .001, d = 0.44, and human nature, $t(284)$ = 2.29, p = .023, d = 0.27.

To evaluate the overall pattern, we conducted an omnibus test using a 2 (communication medium: text or voice) × 2 (agreement: evaluator agreed or disagreed with the communicator) × 2 (measure: human-uniqueness or human-nature traits) mixed-model ANOVA. This analysis revealed significant main effects of communication medium, $F(1, 847)$ = 22.19, p < .001, η_p^2 = .03; agreement, $F(1, 847)$ = 144.07, p < .001, η_p^2 = .15; and measure, $F(1, 847)$ = 23.31, p < .001, η_p^2 = .03. These main effects were qualified by an interaction between measure and communication medium, $F(1, 847)$ = 6.56, p = .011, η_p^2 = .01, indicating that communication medium had a bigger effect on evaluations of human uniqueness than on evaluations of human nature. The remaining interactions were nonsignificant, Fs < 0.31, ps > .250, η_p^2s < .01.

Experiment 4: Mindless Voices

We hypothesized that speech is humanizing because a speaker's voice contains paralinguistic cues that reveal uniquely human mental processes related to thinking and feeling (Schroeder & Epley, 2016). Text lacks these cues, and therefore, evaluators who read, rather than listen to, another person's beliefs may judge that person to be less mentally capable, and hence as having less uniquely human capacities. This reasoning suggests that removing the authentic paralinguistic cues in a person's voice, such as by reducing intonation, may make the person seem less humanlike, as we observed when communicators were evaluated by their text alone. We tested this hypothesis in Experiment 4 by asking participants to listen to transcribed speeches from Experiment 2 that were converted to speech via text-to-speech computer software and therefore delivered by "mindless" voices lacking authentic human paralinguistic cues. We expected that listening to these mindless voices, rather than to the communicators' authentic voices, would result in lower evaluations of communicators' humanlike traits, much as reading communicators' transcribed speeches did. We further predicted that the difference in evaluations of communicators created by listening to mindless voices rather than to communicators' authentic voices would be mediated by differences in paralinguistic cues.

Method

Participants

Communicators. The videotaped speeches and written statements provided by the 8 communicators from Experiment 2 were used in this experiment as well. We used an online text-to-speech program (http://www.fromtexttospeech.com/) to create eight speeches for the mindless-voice condition. The mindless-voice speeches and the authentically voiced speeches therefore contained the same semantic content. To match the communicators' demographics and mimic a typical human speaking pace, in the text-to-speech program we selected U.S. English language, medium speed, and "Alice" and "George" voices for female and male speakers, respectively.

Evaluators. We targeted a sample of 480 evaluators, in an attempt to obtain at least 10 evaluators for each communicator in each experimental condition. Because of a minor error in the survey that required restarting data collection—an incorrect password that appeared at the end of the survey, which made it difficult for participants to receive payment—we collected data from more than the planned number of evaluators. Our final sample consisted of 666 Mechanical Turk workers, who participated in exchange for $0.75 each (mean age = 36.62 years, *SD* = 12.24; 58 female, 40% male, 2% with unreported gender).

Procedure. The evaluators first completed an audio test that ensured they could hear the audio content. They then reported their own political beliefs, rated how favorably they viewed each candidate (0 = *extremely unfavorable*, 6 = *extremely favorable*), listened to or read a communicator's speech; and finally completed the survey used in Experiment 2 to assess evaluations of the communication (without the items measuring persuasion). We randomly assigned each evaluator to 1 of 24 experimental conditions in a 3 (communication medium: authentic voice, mindless voice, or transcript) × 8 (communicator) between-participants design. Some evaluators perceived the assigned communicator's candidate choice favorably, and others perceived the assigned communicator's candidate choice unfavorably, so there were 48 unique experimental conditions. The evaluators in all the conditions learned that the communicators had been asked to state which candidate they supported for president and to explain why they supported this candidate. We additionally told the evaluators in the mindless-voice condition that we had transcribed the communicator's speech and "asked someone to read it aloud." All the evaluators were then told to listen to or read the entire speech carefully. After they had done this, they completed a survey.

We used two methods to ensure that the evaluators read or listened to the entire speech. First, we programmed the survey to automatically pause the audio clip if an evaluator clicked outside of the window containing the audio player. Second, the evaluators could not proceed to the survey for the full duration of the audio clip in the two voice conditions or for at least 20 s in the transcript condition.

We included the same three attention checks as in Experiment 2, with one alteration. The response options for the question "In what form did we show you the participant's opinions?" were changed to "listening to the participant," "listening to someone other than the participant who was reading the participant's transcribed speech," and "reading the participant's transcribed speech." This question allowed us to measure whether participants in the mindless-voice condition understood that they were not listening to the actual communicator.

We expected that communicators would seem more mindless when their beliefs were presented by computer-generated voices, which did not contain authentic paralinguistic cues to thought or feeling. As a manipulation check, we asked the evaluators in the voice conditions (n = 406) five questions concerning their thoughts about the speaker (in randomized order): (a) "How authentic did the speaker sound?" (b) "How genuine did the speaker sound?" (c) "How objective did the speaker sound?" (reverse-scored) (d) "How passionate did the speaker sound?" and (e) "How humanlike did the speaker sound?" (0 = *not at all*, 6 = *very*). We then asked the evaluators to complete the same humanization scale used in Experiments 1 through 3 (α = .86 for human uniqueness and .85 for human nature), plus other measures described in the Supplemental Material.

Results

Thirty-four participants failed one or more of our three attention checks, which left a final sample of 632 evaluators. We excluded more participants in the mindless-voice condition (n = 20) than in the other conditions (authentic-voice: n = 6; transcript: n = 8), $\chi^2(2, N = 666) = 10.10$, p = .006, in part because they were more likely to mistakenly report that they had read a transcript.

We coded each evaluator's agreement with the assigned communicator using the same method as in Experiments 1 through 3. Unexpectedly, evaluators' agreement with communicators varied by communication medium, $\chi^2(2, N = 632) = 10.46$, p = .005; more evaluators agreed with communicators in the transcript condition (40.3%) than in the mindless-voice condition (26.8%) or the authentic-voice condition (28.9%). This difference does not affect the interpretation of our results for two reasons. First, communication-medium condition cannot have affected agreement because we asked the evaluators about the candidate they supported before they were assigned to a communication medium condition. The difference in agreement between conditions therefore appears to be a failure of random assignment. Second, we analyzed the effect of communication medium separately for participants who disagreed and who agreed with the assigned communicator.

Table 1 summarizes the ratings of human uniqueness and human nature in the three communication-medium

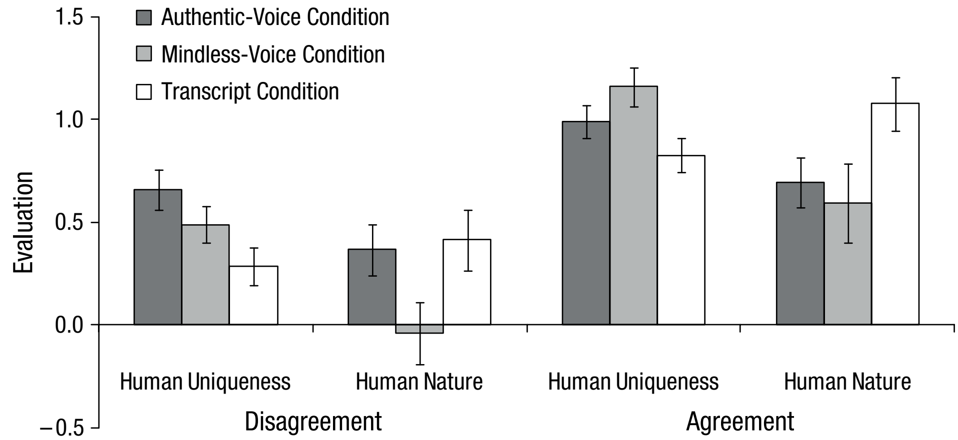

Fig. 4. Evaluations of communicators' human-uniqueness and human-nature traits in the authentic-voice, mindless-voice, and transcript conditions of Experiment 4. Results are presented separately for evaluators who agreed and who disagreed with the communicators they rated. Error bars represent ±1 *SEM*.

conditions, separately for evaluators who agreed and who disagreed with their assigned communicators. As in Experiments 1 through 3, evaluators dehumanized a communicator with an opposing viewpoint less when they heard what the person had to say than when they read it. Communication medium affected evaluations of both the communicators' human uniqueness, $F(2, 425) = 3.96$, $p = .020$, $\eta^2 = .02$, and their human nature, $F(2, 425) = 8.18$, $p < .001$, $\eta_p^2 = .04$ (see Fig. 4). The results of the prior experiments were replicated: Evaluators in the authentic-voice condition rated a communicator who disagreed with them more highly on human uniqueness ($M = 0.66$, $SD = 1.11$, 95% CI = [0.48, 0.84]) than did evaluators in the transcript condition ($M = 0.28$, $SD = 1.14$, 95% CI = [0.09, 0.48]), $t(425) = 2.81$, $p = .005$, $d = 0.34$. However, evaluators in the authentic-voice condition did not rate a communicator who disagreed with them more highly on human nature ($M = 0.37$, $SD = 1.00$, 95% CI = [0.20, 0.53]) than did evaluators in the transcript condition ($M = 0.41$, $SD = 0.94$, 95% CI = [0.25, 0.57]), $t(425) = -0.37$, $p > .250$, $d = 0.04$. We note that across our experiments, the effect of voice on judgments of human nature was consistently weaker than the effect of voice on judgments of human uniqueness, which perhaps suggests that voice conveys mental capacities related to cognition and thinking more clearly than capacities related to emotional experience and feeling.

Our primary interest in this experiment was testing whether evaluators were more likely to dehumanize communicators when we removed authentic paralinguistic cues from their voices. Specifically, we predicted that among evaluators who disagreed with communicators, evaluators in the mindless-voice condition would dehumanize communicators more than those in authentic-voice condition. As shown in Figure 4, ratings of human uniqueness in the mindless-voice condition fell in between ratings of human uniqueness in the authentic-voice and transcript conditions, $t(425) = 1.31$, $p = .190$, $d = 0.15$, and $t(425) = -1.55$, $p = .122$, $d = 0.18$, respectively. The lack of a significant difference between this condition and the others did not fully support our hypothesis. However, evaluators' ratings of the communicators' human nature was significantly lower in the mindless-voice condition, than in both the authentic-voice condition, $t(425) = -3.31$, $p = .001$, $d = -0.39$, and the transcript condition, $t(425) = -3.63$, $p < .001$, $d = -0.43$. In a post hoc analysis, we combined the human-uniqueness and human-nature items ($\alpha = .91$; all items loaded onto one factor accounting for 49.17% of the variance, with factor loadings > .58) and tested the effect of voice type on this overall measure of humanization. In this analysis, evaluators dehumanized communicators more in the mindless-voice condition ($M = 0.35$, $SD = 0.99$, 95% CI = [0.18, 0.52]) than in the authentic-voice condition ($M = 0.51$, $SD = 0.96$, 95% CI = [0.35, 0.67]), $t(425) = 2.49$, $p = .013$, $d = 0.29$.

When the evaluators agreed with the communicators, we observed no effect of communication medium on ratings of human uniqueness, $F(2, 201) = 1.47$, $p = .232$, $\eta^2 = .01$, but did find an unpredicted effect of communication medium on ratings of human nature, $F(2, 201) = 3.38$, $p = .036$, $\eta^2 = .03$. This effect was driven by evaluators in the transcript condition, who rated communicators more highly on human nature ($M =$

1.07, SD = 1.16, 95% CI = [0.83, 1.32]) than did evaluators in the mindless-voice condition (M = 0.59, SD = 1.43, 95% CI = [0.21, 0.98]), $t(201)$ = 2.37, p = .019, d = 0.40, and evaluators in the authentic-voice condition (M = 0.69, SD = 1.00, 95% CI = [0.43, 0.96]), $t(201)$ = 1.90, p = .059, d = 0.32. Evaluations in cases of agreement were inconsistent across our experiments, and because this unpredicted effect did not emerge in any of the other experiments, we do not discuss it further.

To evaluate the overall pattern, we conducted an omnibus test using a 3 (communication medium: authentic voice, mindless voice, or transcript) × 2 (agreement: evaluator agreed or disagreed with the communicator) × 2 (measure: human-uniqueness or human-nature traits) mixed-model ANOVA. This analysis revealed significant main effects of agreement, $F(1, 626)$ = 34.82, p < .001, η_p^2 = .05, and measure, $F(1, 626)$ = 38.26, p < .001, η_p^2 = .05. The effect of measure was qualified by an interaction between measure and communication medium, $F(2, 626)$ = 38.26, p < .001, η_p^2 = .11. All other interactions and the remaining main effect were nonsignificant, Fs < 1.20, ps > .250, η_p^2s < .01. The interaction between measure and communication medium indicated that communication medium again influenced evaluations of human uniqueness more than evaluations of human nature.

Differences between authentic voices and mindless voices. Our manipulation check (five items; α = .85) confirmed that the evaluators believed that the communicators' authentic voices sounded more authentic (M = 4.30, SD = 1.08) than the mindless voices (M = 2.12, SD = 1.51), $t(401)$ = 16.72, p < .001, d = 1.67. To examine which paralinguistic cues mediated the effect of voice on evaluations of the communicators' humanlike traits, we first used Praat software (Boersma & Weenink, 2016) to extract paralinguistic cues commonly studied in nonverbal psychology (Hughes, Mogilski, & Harrison, 2014; Laplante & Ambady, 2003): mean pitch, intonation (standard deviation of pitch), speech length (to estimate pace of speaking), and mean percentage of pauses. We followed standard procedure to extract these cues from each of the 16 voice clips (8 speaker voices and 8 mindless voices).[2] On the basis of prior research, we predicted that mental capacities would be related to variability in paralinguistic cues, especially to variance in pitch and pace (i.e., intonation and pauses; Banziger & Scherer, 2005; Schroeder & Epley, 2016).

The mindless voices differed from the communicators' authentic voices on all of the paralinguistic cues we measured. Compared with the mindless voices, communicators' authentic voices had marginally higher mean pitch (authentic voices: M = 145.32 Hz, SD = 51.34; mindless voices: M = 122.03 Hz, SD = 33.56), paired-samples $t(7)$ = 2.16, p = .067, 95% CI for the mean difference = [−2.19, 48.75], d = 0.54, and more intonation (authentic voices: M = 50.95, SD = 24.31; mindless voices: M = 23.30, SD = 5.16), paired-samples $t(7)$ = 3.27, p = .014, 95% CI for the mean difference = [7.65, 47.66], d = 1.57. Also, total speech length was greater for the mindless voices (M = 114.24 s, SD = 50.87) than for the authentic voices (M = 85.27 s, SD = 13.89), an indication of a slower speaking pace, paired-samples $t(7)$ = 4.51, p = .003, 95% CI for the mean difference = [13.78, 44.15], d = 0.78, and there was a higher percentage of pauses for the authentic voices (M = 61.00%, SD = 10.84) than for the mindless voices (M = 37.63%, SD = 9.54), paired-samples $t(7)$ = 5.19, p = .001, 95% CI for the mean difference = [12.72, 34.03], d = 2.29.

Mediation by paralinguistic cues. To determine which, if any, paralinguistic cues mediated the effect of communication-medium condition on humanization (ratings of human uniqueness and human nature combined into a single composite, α = .91), we conducted a series of multilevel regression models and computed Sobel tests to estimate the indirect effect for each cue. We used this method because our potential mediators were computed from the communicators, but the humanness ratings were collected from the evaluators; therefore, we had to conduct hierarchical analyses, which are not well suited for bootstrapped mediation models. We tested each cue separately because the cues were highly intercorrelated (for mean pitch, intonation, and pauses, rs ranged from .82 to .95). We were therefore unable to meaningfully calculate the unique influence of any one paralinguistic cue controlling for the others. In these analyses, we included data from the evaluators who were in the authentic-voice and mindless-voice conditions and who disagreed with their communicators' choice of presidential candidate (n = 293).

We first tested the direct effect of voice type on humanization by conducting a multilevel regression model on evaluations with the single predictor of voice type, controlling for speaker fixed effects. The direct effect was significant, β = 0.33, SE = 0.11, p = .003. We next tested whether voice type predicted paralinguistic cues. Results were consistent with the t tests comparing paralinguistic cues in the mindless and authentic voices. Voice type significantly predicted all of the paralinguistic cues, tested in separate regression models—mean pitch: β = 22.24, SE = 1.67, p < .001; intonation: β = 31.93, SE = 1.33, p < .001; speech length: β = 28.05, SE = 0.99, p < .001; and percentage of pauses: β = 0.27, SE = 0.004, p < .001. We then tested which of these paralinguistic cues significantly predicted ratings of humanness and found that only intonation and percentage of pauses did so, β = 0.01, SE = 0.003, p = .009, and

$\beta = 0.92$, $SE = 0.37$, $p = .014$, respectively. In contrast, mean pitch marginally predicted and speech length nonsignificantly predicted ratings of humanness, $\beta = 0.005$, $SE = 0.002$, $p = .059$, and $\beta = 0.002$, $SE = 0.002$, $p = .344$, respectively (Preacher & Hayes, 2008).

Finally, we tested whether including mean pitch, intonation, speech length, or percentage of pauses in our models would remove the effect of voice type on evaluations. When we included mean pitch and speech length in two separate regression models, voice type remained a significant predictor of ratings of humanness, $\beta = 0.32$, $SE = 0.11$, $p = .005$, and $\beta = 0.38$, $SE = 0.11$, $p = .001$, respectively. But including intonation or pause percentage (in two separate models) made the effect nonsignificant, $\beta = 0.23$, $SE = 0.16$, $p = .143$, and $\beta = 0.36$, $SE = 0.21$, $p = .116$, respectively. Sobel tests for each of these cues indicated that the indirect effects were statistically significant in the case of mean pitch ($z = 2.46$, $p = .014$), intonation ($z = 3.30$, $p = .001$), and percentage of pauses ($z = 2.48$, $p = .013$), but not speech length ($z = 1.00$, $p = .318$). These results suggest that intonation and the percentage of pauses are the most viable mediators and can at least partly account for the effect of voice on humanization. Variability in a person's voice (in this case, intonation and pauses) may communicate the presence of a humanlike mind.

General Discussion

When two people hold different beliefs, there is a tendency not only to recognize a difference of opinion but also to denigrate the mind of one's opposition. The "other side" in political disputes is seen not simply as thinking differently about a topic, but also as being less capable of thinking altogether (Ross & Ward, 1996). Denigrating the mind of another person is the essence of dehumanization—seeing that person as less capable of thinking or feeling than oneself, as more like a nonhuman animal than like a mentally sophisticated human being.

These four experiments demonstrated that the medium of communication may moderate the tendency to dehumanize the opposition. Because another person's mind cannot be experienced directly, its quality must be inferred from indirect cues. The human voice contains paralinguistic cues that reveal underlying mental processing involved in thinking and feeling. These cues are absent from text-based media, and as a result, individuals from the opposition seem to have more uniquely human capacities when people *hear* what they have to say than when they *read* similar content. Inferences about the humanlike capacities of other people may depend on the medium through which they communicate.

In our experiments, participants evaluated communicators who agreed or disagreed with them on polarizing issues (Experiment 1) or political preferences (Experiments 2–4). We observed an inconsistent influence of communication-medium condition in cases of agreement, but in cases of disagreement, we observed a reliable tendency for communicator to be dehumanized less when evaluators heard their voices than when evaluators read the same content. This effect occurred both when the semantic content was presented in the form of speech transcriptions and when it was presented via the communicators' own written statements. Adding visual cues to a communicator's voice did not systematically increase evaluations of the communicator's mental capacities. This finding suggests either that humanizing cues are unique to the voice or that such cues are redundant in visual and vocal media. Our experiments did not directly compare the humanizing capacities of visual and vocal cues. Instead, our data demonstrate that removing voice (via text), or altering paralinguistic cues in a person's authentic voice (e.g., by using a computer-generated voice), can result in dehumanization. Our findings further suggest that reliable individual differences in people's voices may be related to humanization; for instance, individuals with voices that lack authentic intonation (e.g., monotone voices) may be perceived as less humanlike than others.

In each experiment, we observed a stronger effect on evaluations of human uniqueness (traits related to reasoning and cognition) than on evaluations of human nature (traits related to emotional experience and interpersonal warmth). This finding was unpredicted, and its meaning is unclear. It could reflect a general tendency for voice to convey capacities related to thinking more clearly than capacities related to feeling, or it could reflect the experimental context, in which people communicated their thoughts on an important issue, rather than their emotions or interpersonal experiences. Future research will need to clarify the meaning of this result.

On a theoretical level, our findings integrate research on language and humanization, suggesting that the medium of communication meaningfully influences judgments of uniquely human mental capacities during disagreement. Whereas existing research demonstrates that cues in speech increase accurate understanding of mental states (Epley & Kruger, 2005; Hall & Schmid Mast, 2007; Kruger et al., 2005; Zaki, Bolger, & Ochsner, 2009), our experiments demonstrate that a person's voice reveals something more fundamental: the presence of a humanlike mind capable of thinking and feeling. This research also suggests a new interpersonal determinant of dehumanization. Existing research has focused primarily on intergroup mechanisms, such as when members of one group dehumanize other negatively stereotyped groups (Harris & Fiske, 2009; Vaes, Leyens, Paladino, & Miranda, 2012). Understanding the interpersonal mechanisms that guide dehumanization will suggest novel interventions for changing intergroup relations.

On a practical level, our work suggests that giving the opposition a voice, not just figuratively in terms of language, but also literally in terms of an actual human voice, may enable partisans to recognize a difference in beliefs between two minds without denigrating the minds of the opposition. Modern technology is rapidly changing the media through which people interact, enabling interactions between people around the globe and across ideological divides who might otherwise never interact. These interactions, however, are increasingly taking place over text-based media that may not be optimally designed to achieve a user's goals. Individuals should choose the context of their interactions wisely. If mutual appreciation and understanding of the mind of another person is the goal of social interaction, then it may be best for the person's voice to be heard.

Action Editor

Jamin Halberstadt served as action editor for this article.

Author Contributions

J. Schroeder and N. Epley conceived and designed the experiments. J. Schroeder and M. Kardas collected the stimuli and data for the experiments and performed the data analysis. J. Schroeder and N. Epley wrote the manuscript; M. Kardas provided revisions. All the authors approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617713798

Open Practices

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/nm8vf/. All stimuli that the communicators generated for the experiments reported in this article are available upon request from the authors. The design and analysis plans for the experiments were preregistered at the Open Science Framework and can be accessed at https://osf.io/nm8vf. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617713798. This article has received badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

Notes

1. To ensure that the communicators were informed enough about the election to discuss it, we asked them two more questions before they wrote and spoke about their opinions. First, we asked, "How closely have you followed this year's Presidential election compared to other people?" Second, we asked, "How informed do you feel about this year's election compared to other people?" The communicators responded to both questions on 9-point scales (−4 = *much less closely/informed than other people*, 0 = *no more or less closely/informed than other people*, 4 = *much more closely/informed than other people*). Results revealed that the communicators felt relatively well informed, responding above the midpoint for each scale, $t(9)\text{s} > 2.33$, $p\text{s} < .045$, $d\text{s} > 0.74$.

2. To compute each speaker's pitch profile in Praat, we first set a fixed time step of 0.01 s. We set the pitch range to 150 to 500 Hz for female speakers and 75 to 500 Hz for male speakers. We used the autocorrelation analysis method in the pitch settings. To export the pitch, we selected the entire pitch profile and saved the pitch listing as a text file that we imported into Excel. We then computed the average and standard deviation of the pitch in Excel. The number of seconds covered in these pitch profiles composed our measure of speech length. To compute the percentage of pauses, we counted the number of blank cells in the pitch profile (each Excel cell represented 0.01 s) and divided that number by the speech's duration (i.e., the total number of cells in the speech).

References

Ames, D. R. (2004). Strategies for social inference: A similarity contingency model of projection and stereotyping in attribute prevalence estimates. *Journal of Personality and Social Psychology*, *87*, 573–585.

Banziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication*, *46*, 252–267.

Bastian, B., & Haslam, N. (2010). Excluded from humanity: The dehumanizing effects of social ostracism. *Journal of Experimental Social Psychology*, *46*, 107–113.

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer (Version 6.0.30) [Computer software]. Retrieved from http://www.praat.org/

Carlin, G. (1984). *Carlin on Campus* [Record]. United States: Eardrum Records.

Epley, N., & Kruger, J. (2005). When what you type isn't what they read: The perseverance of stereotypes and expectancies over e-mail. *Journal of Experimental Social Psychology*, *41*, 414–422.

Epley, N., & Waytz, A. (2010). Mind perception. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The handbook of social psychology* (5th ed., pp. 498–541). New York, NY: Wiley.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*, 619.

Hall, J. A., & Schmid Mast, M. (2007). Sources of accuracy in the empathic accuracy paradigm. *Emotion*, *7*, 438–446.

Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology*, *20*, 192–231.

Haslam, N. (2006). Dehumanization: An integrated review. *Personality and Social Psychology Review*, *10*, 252–264.

Haslam, N., & Bain, P. (2007). Humanizing the self: Moderators of the attribution of lesser humanness to others. *Personality and Social Psychology Bulletin*, *33*, 57–68.

Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you: Attributing humanness to self and others. *Journal of Personality and Social Psychology*, *89*, 937–950.

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, *65*, 399–423.

Haslam, N., Loughnan, S., & Holland, E. (2013). The psychology of humanness. In S. J. Gervais (Ed.), *Nebraska Symposium on Motivation: Vol. 60. Objectification and (de)humanization* (pp. 25–52). New York, NY: Springer Science.

Hughes, S. M., Mogilski, J. K., & Harrison, M. A. (2014). The perception and parameters of intentional voice manipulation. *Journal of Nonverbal Behavior*, *38*, 107–127.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.

Kennedy, K. A., & Pronin, E. (2008). When disagreement gets ugly: Perceptions of bias and the escalation of conflict. *Personality and Social Psychology Bulletin*, *34*, 833–848.

Krueger, J. (2000). The projective perception of the social world: A building block of social comparison processes. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 323–351). New York, NY: Plenum/Kluwer.

Kruger, J., Epley, N., Parker, J., & Ng, Z. (2005). Egocentrism over e-mail: Can people communicate as well as they think? *Journal of Personality and Social Psychology*, *89*, 925–936.

Laplante, D., & Ambady, N. (2003). On how things are said: Voice tone, voice intensity, verbal content, and perceptions of politeness. *Journal of Language and Social Psychology*, *22*, 434–441.

Leyens, J.-P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., & Gaunt, R. (2000). The emotional side of prejudice: The role of secondary emotions. *Personality and Social Psychology Review*, *4*, 186–197.

Malle, B. F., Knobe, J., & Nelson, S. (2007). Actor-observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology*, *93*, 491–514.

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say 'hello'? Personality impressions from brief novel voices. *PLoS ONE*, *9*(3), Article e90779. doi:10.1371/journal.pone.0090779

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*, 707–727.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891.

Pronin, E. (2009). The introspection illusion. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 41, pp. 1–67). Burlington, MA: Academic Press.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, *28*, 369–381.

Ross, L., & Ward, A. (1996). Naïve realism in everyday life: Implications for social conflict and misunderstanding. In T. Brown, E. Reed, & E. Turiel (Eds.), *Values and knowledge* (pp. 103–135). Hillsdale, NJ: Erlbaum.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, *32*, 76–92.

Schroeder, J., & Epley, N. (2015). The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal. *Psychological Science*, *26*, 877–891.

Schroeder, J., & Epley, N. (2016). Mistaking minds and machines: How speech affects dehumanization and anthropomorphism. *Journal of Experimental Psychology: General*, *145*, 1427–1437.

Vaes, J., Leyens, J.-P., Paladino, M. P., & Miranda, M. P. (2012). We are human, they are not: Driving forces behind outgroup dehumanisation and the humanisation of the ingroup. *European Review of Social Psychology*, *23*, 64–106.

Waytz, A., Schroeder, J., & Epley, N. (2014). The lesser minds problem. In P. Bain, J. Vaes, & J. P. Leyens (Eds.), *Humanness and dehumanization* (pp. 49–67). London, England: Psychology Press.

Wegner, D. M., & Gilbert, D. T. (2000). Social psychology—the science of human experience. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 1–9). Philadelphia, PA: Psychology Press.

Weisbuch, M., Pauker, K., & Ambady, N. (2009). The subtle transmission of race bias via televised nonverbal behavior. *Science*, *326*, 1711–1714.

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, *25*, 1115–1125.

Zaki, J., Bolger, N., & Ochsner, K. (2009). Unpacking the informational bases of empathic accuracy. *Emotion*, *9*, 478–487.