Demographics and (Equal?) Voice: Assessing Participation in Online Deliberative Sessions
Supporting Information

October 17, 2019

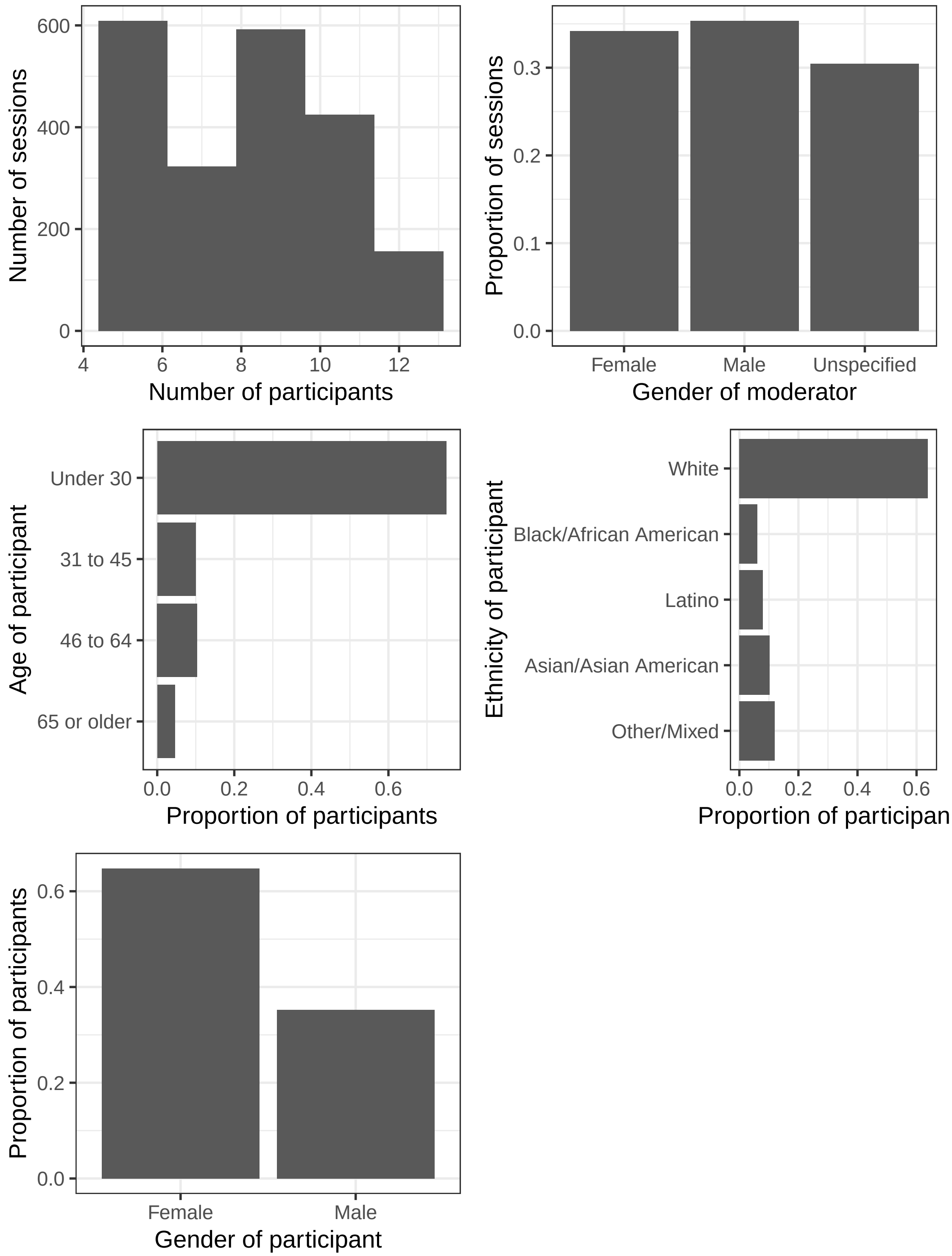
Contents

A1 Descriptive Statistics

Table A1: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Scale(ln(number of chats))	2105	0.478	0.759	−1.471	0.071	0.999	2.387
Scale(ln(number of words))	2105	0.488	0.664	−1.618	0.312	0.905	1.489
Satisfaction with common ground	1085	4.147	0.874	1.000	4.000	5.000	5.000
Female	1720	0.648	0.478	0.000	0.000	1.000	1.000
Age 31 to 45	1670	0.132	0.338	0.000	0.000	0.000	1.000
Age 46 to 64	1670	0.168	0.374	0.000	0.000	0.000	1.000
Age 65 or older	1670	0.046	0.210	0.000	0.000	0.000	1.000
Non-white	1631	0.362	0.481	0.000	0.000	1.000	1.000
Pre-session preference deviation	2105	0.459	0.129	0.000	0.375	0.532	1.194
Pre-session preference correlation	2105	0.197	0.215	−1.000	0.082	0.329	0.866
Scale(number of participants	2105	0.464	0.676	−0.530	−0.206	1.090	1.738
Female moderator	2105	0.342	0.475	0	0	1	1
Unspecified moderator	2105	0.305	0.460	0	0	1	1
Scale(proportion female)	2074	0.053	0.771	−2.368	−0.529	0.574	1.310
Proportion of common ground participant supported	1905	0.471	0.300	0.000	0.250	0.667	1.000
Proportion of words from female participants	2105	0.213	0.321	0	0	0.5	1

Figure A1: Demographic Characteristics of Sessions.



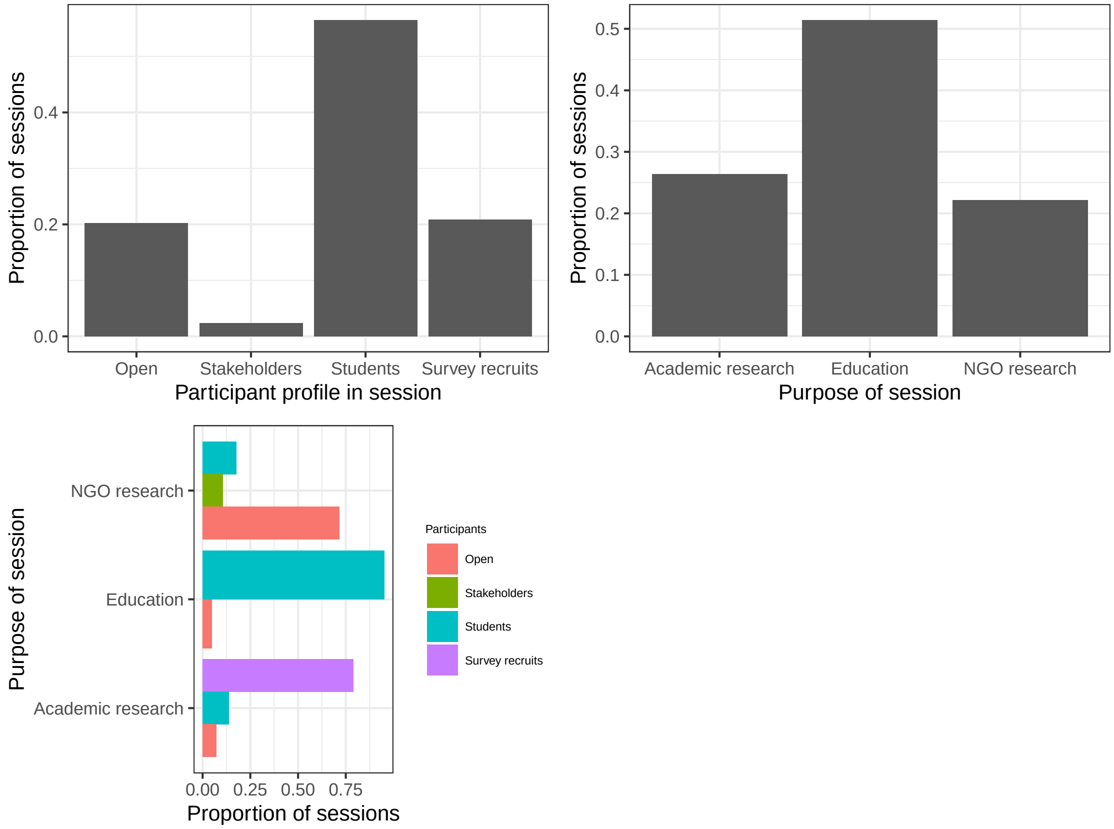
A2 Description of Sessions

As noted in the main paper, the makeup of the sessions and the purposes for which they were conducted vary between forums. Figure A2 shows a breakdown along both of these dimensions. The top-left chart shows the proportion of sessions that involve four different types of participants. ”Open” sessions are those in which the researcher recruited from an online group, including those who signed up to participate through NIFI’s site for the CGAs. ”Stakeholds” were sessions where the research group attempted to recruit those who were affected by the particular issue, usually within a particular locality, to participate in the discussion. ”Students” were recruited from universities to discuss issues. ”Survey recruits” were recruited to participate by a professional survey firm (Qualtrics and Naviscent) to participate in the sessions. As the chart shows, a little over half of the participants were students.

The top-right chart shows the breakdown by the purpose of the study. ”Academic research” includes situations where the researcher was conducting the session in order to test the efficacy of the CGA platform or to perform and experiment about the impact of deliberation. These could include any of the three populations, although most of them came from the studies that used professional survey firms for recruiting, as shown in the bottom chart. ”Education” denotes situations where the CGA was used as part of a teaching curriculum, either about deliberation or current events. There is some overlap between these categories – some professors incorporated it into their classroom, but might have used the results in a subsequent study. Where there was uncertainty about these categories, the transcripts were reviewed to check whether the respondents appeared to be part of a class (i.e., there was some mention of a curriculum). As the bottom chart shows, almost all of these participants were students, with the exception of one session where it appeared that a session was conducted by an education team with an open recruiting system as a way for one of their moderators to get practice with the system, which they were subsequently going to use for a class. While this session was conducted with an open population, and was conducted in the same manner as other open sessions, we coded its purpose as primarily educational. Finally ”NGO research” denotes sessions conducted by Kettering, the American Library Association, or another similar group. These were predominately drawn from open groups, but with some sessions done in cooperation with universities among student populations or with recruited stakeholders. Since the purpose of these sessions was primarily for research and reports by the NGO, they fell into this category.

There are likely to be some differences between groups. In particular, those groups that are made up of students are more likely to know each other outside of the session, and we did have concern that this might affect the results. In the main paper, we enter the population categories as fixed effects, with ”open” sessions as the baseline. We also tested the interaction with gender to see if this changed our results at all. We found no statistically significant difference among the types of sessions, suggesting that the main results were produced by the technology used, rather than the population from which the participants were drawn.

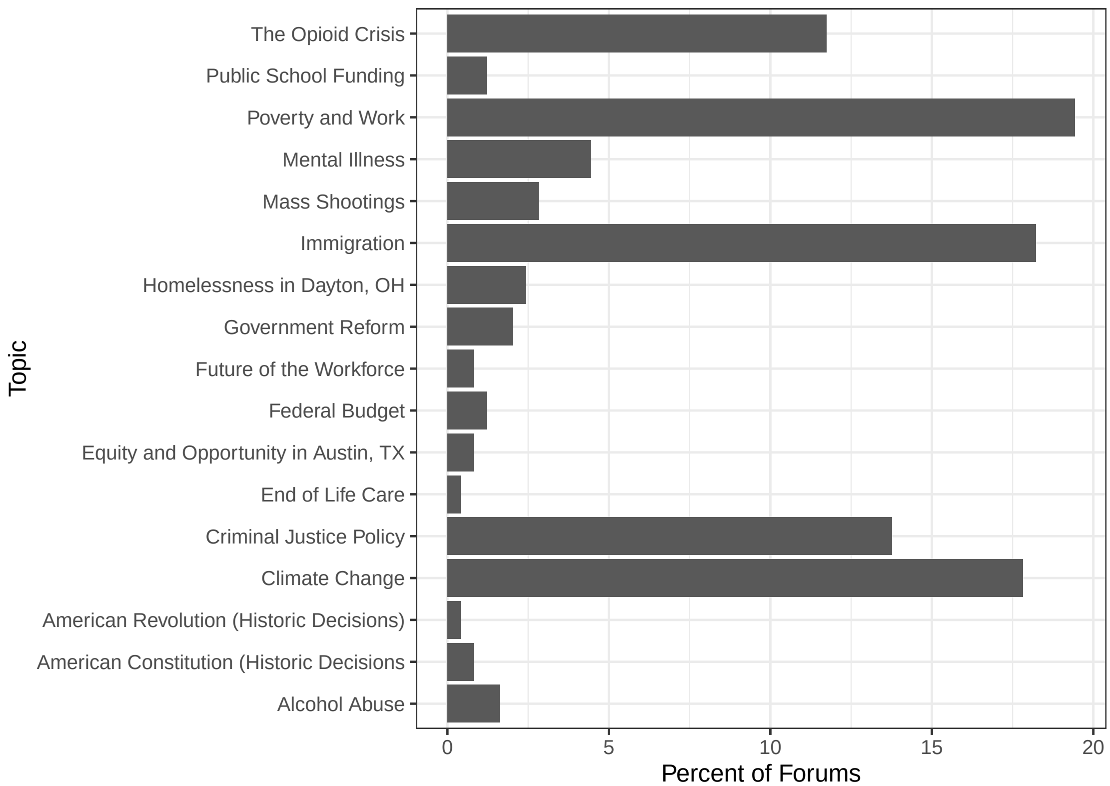
Figure A2: Characteristics of Session Populations and Purposes.



A3 Topics of Sessions

As noted in the main paper, the sessions also varied in the topic they covered. As noted in the main paper, the vast majority of these were on topics where we think the dictionaries of Karpowitz and Mendelberg (2014) will work relatively well. Figure A3, shows the proportion of session conducted under each topic. The vast majority of sessions were conducted with one of five topics: the opioid crisis, poverty and work, immigration, criminal justice policy, and climate change. The rest were conducted under a variety of other topics.

Figure A3: Proportion of Sessions Conducted Under Each Topic.



A4 Full Results

This section shows the full tables for the plots in the main paper. Table A2 shows the results underlying Figures 2 and 3. Table A3 shows the results underlying Figure 6. The table notes indicate how variables have been scaled for convergence and/or transformed to accounted for skewed distributions. Substantive effects reported in the text of the main paper reverse the scaling and transformations. Scaling was done as $(x - \bar{x})/sd$, such that a 1 point change in the scaled variable indicates a 1 standard deviation change. So, for example, for a variable that has been scaled, a 1-point change in the regression coefficient indicates a 1-standard deviation change. The substantive result is then calculated as $1 * sd(x)$. If that variable is also logged, it would be $exp(1 * sd(x))$. Logged dependent variables have a somewhat different interpretation, since they are based on the geometric rather than the arithmetic mean. So, for example, in Model 1 of Table A2, the coefficient for female participants is -0.077 and the standard deviation of natural log of the number of chats is 1.024. Thus, the difference between men and women would be $exp(sd(x) * -0.077) = 0.924$, meaning that women had about 0.924 the proportion of chats as men, or about 8% fewer posted chats.

Table A2: The Effect of Participant Characteristics on the log Number of Chats and Words in CGA Sessions

	Chats	**Words**	**Chats**	**Words**	**Chats**	**Words**	**Chats**	**Words**
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Individual Level								
Female	−0.078***	−0.034	−0.073**	−0.033	−0.126***	−0.067**	−0.028	0.002
	(0.029)	(0.026)	(0.030)	(0.026)	(0.037)	(0.032)	(0.062)	(0.054)
31 to 45	0.157***	0.075*	0.157***	0.075*	0.152***	0.072	0.157***	0.074*
	(0.050)	(0.043)	(0.050)	(0.044)	(0.050)	(0.044)	(0.051)	(0.044)
46 to 64	0.131**	0.020	0.132**	0.020	0.127**	0.018	0.126**	0.018
	(0.051)	(0.044)	(0.052)	(0.044)	(0.051)	(0.044)	(0.051)	(0.044)
65 or older	−0.167**	−0.270***	−0.166**	−0.270***	−0.174**	−0.275***	−0.164**	−0.269***
	(0.076)	(0.065)	(0.076)	(0.065)	(0.076)	(0.065)	(0.076)	(0.065)
Non-white	−0.179***	−0.184***	−0.178***	−0.184***	−0.180***	−0.185***	−0.182***	−0.186***
	(0.030)	(0.026)	(0.030)	(0.026)	(0.030)	(0.026)	(0.030)	(0.026)
Preference deviation	−0.028	−0.014	−0.030	−0.015	−0.021	−0.009	−0.035	−0.018
	(0.112)	(0.097)	(0.112)	(0.097)	(0.112)	(0.097)	(0.112)	(0.097)
Forum Level								
N participants	−0.052	−0.034	−0.052	−0.034	−0.056	−0.036	−0.055	−0.035
	(0.037)	(0.029)	(0.037)	(0.029)	(0.037)	(0.029)	(0.037)	(0.029)
Female moderator	−0.044	−0.032	−0.045	−0.032	−0.123	−0.085	−0.042	−0.030
	(0.066)	(0.053)	(0.066)	(0.053)	(0.075)	(0.061)	(0.066)	(0.053)
Unspecified moderator	−0.529***	−0.302***	−0.529***	−0.302***	−0.523***	−0.299***	−0.523***	−0.299***
	(0.073)	(0.059)	(0.074)	(0.059)	(0.073)	(0.059)	(0.073)	(0.059)
Proportion female	−0.034	−0.041	−0.053	−0.045	−0.034	−0.040	−0.030	−0.038
	(0.033)	(0.026)	(0.044)	(0.036)	(0.033)	(0.026)	(0.033)	(0.026)
Education	−0.137*	−0.091	−0.139**	−0.092	−0.135*	−0.090	−0.067	−0.047
	(0.071)	(0.057)	(0.071)	(0.057)	(0.071)	(0.057)	(0.085)	(0.071)
Stakeholders	−0.219	−0.124	−0.219	−0.124	−0.213	−0.121	−0.236	−0.132
	(0.175)	(0.164)	(0.176)	(0.164)	(0.175)	(0.163)	(0.218)	(0.200)
Survey	−0.035	0.039	−0.036	0.039	−0.031	0.043	−0.049	0.047
	(0.093)	(0.079)	(0.093)	(0.079)	(0.093)	(0.079)	(0.108)	(0.093)
Cross-Level Interactions								
Female x proportion female			0.028	0.006				
			(0.044)	(0.038)				
Female x female moderator					0.129**	0.086*		
					(0.058)	(0.051)		
Female x education							−0.107	−0.066
							(0.073)	(0.064)
Female x stakeholders							0.023	0.011
							(0.186)	(0.162)
Female x survey							0.030	−0.007
							(0.086)	(0.075)
Random Effects								
# of Forums	275	275	275	275	275	275	275	275
Forum Standard Deviation	0.307	0.224	0.307	0.224	0.305	0.223	0.305	0.223
# of Topics	20	20	20	20	20	20	20	20
Topic Standard Deviation	0.072	0.104	0.072	0.104	0.072	0.103	0.068	0.103
Constant	0.952***	0.820***	0.947***	0.819***	0.980***	0.840***	0.923***	0.797***
	(0.101)	(0.087)	(0.101)	(0.087)	(0.101)	(0.088)	(0.106)	(0.092)
N	1609	1609	1609	1609	1609	1609	1609	1609
Log Likelihood	−1347.623	−1111.592	−1349.622	−1113.933	−1347.102	−1112.213	−1349.487	−1115.441
AIC	2729.245	2257.183	2735.244	2263.865	2730.204	2260.427	2738.974	2270.882
BIC	2820.762	2348.701	2832.145	2360.766	2827.104	2357.328	2846.642	2378.549

***p < .01; **p < .05; *p < .1

Values in table are regression coefficients with standard errors in parentheses. Dependent variables have been scaled to aid in model convergence, so magnitude of coefficients is interpreted in standard deviations. The number of chats and number of words are also transformed by the natural log. The number of participants and proportion of female participants were also scaled. All models are multilevel models with random intercepts by forum and topic.

Table A3: The Effect of Participant Characteristics on Satisfaction with Common Ground

	Satisfaction			
	Model 1	**Model 2**	**Model 3**	**Model 4**
Individual Level				
Female	0.089	0.077	0.038	−0.037
	(0.064)	(0.065)	(0.078)	(0.081)
31 to 45	0.145	0.143	0.145	0.135
	(0.097)	(0.097)	(0.097)	(0.096)
46 to 64	0.147	0.145	0.146	0.149
	(0.094)	(0.094)	(0.094)	(0.093)
65 or older	0.146	0.143	0.138	0.132
	(0.177)	(0.177)	(0.177)	(0.177)
Non-white	−0.130**	−0.133**	−0.130**	−0.139**
	(0.064)	(0.064)	(0.064)	(0.064)
Proportion of common ground participant supports	0.208*	0.206*	0.202*	0.221**
	(0.107)	(0.107)	(0.107)	(0.107)
Forum Level				
Female proportion words	−0.032	−0.034	−0.032	−0.361**
	(0.098)	(0.099)	(0.098)	(0.164)
N participants	−0.076*	−0.077*	−0.080*	−0.075*
	(0.046)	(0.046)	(0.046)	(0.046)
Starting preference correlation	0.188	0.190	0.187	0.188
	(0.143)	(0.143)	(0.142)	(0.142)
Education	−0.039	−0.033	−0.037	−0.051
	(0.092)	(0.092)	(0.091)	(0.091)
Stakeholders	−0.106	−0.104	−0.116	−0.133
	(0.194)	(0.195)	(0.194)	(0.194)
Survey	−0.128	−0.125	−0.122	−0.139
	(0.107)	(0.107)	(0.107)	(0.106)
Female moderator	0.032	0.033	−0.048	0.031
	(0.078)	(0.078)	(0.104)	(0.077)
Unspecified moderator	−0.096	−0.098	−0.091	−0.090
	(0.096)	(0.096)	(0.096)	(0.096)
Proportion female	0.0002	0.040	0.0004	0.009
	(0.046)	(0.069)	(0.046)	(0.046)
Cross-Level Interactions				
Female x proportion female		−0.067		
		(0.085)		
Female x female moderator			0.142	
			(0.122)	
Female x female proportion words				0.468**
				(0.187)
Random Effects				
# of Forums	238	238	238	238
Forum Standard Deviation	0.123	0.126	0.121	0.123
Constant	4.101***	4.118***	4.132***	4.190***
	(0.139)	(0.141)	(0.141)	(0.143)
N	873	873	873	873
Log Likelihood	−1104.224	−1105.462	−1104.735	−1101.852
AIC	2244.448	2248.924	2247.469	2241.704
BIC	2330.343	2339.590	2338.136	2332.371

***p < .01; **p < .05; *p < .1

Values in table are regression coefficients with standard errors in parentheses. The number of participants and proportion of female participants were scaled. All models are multilevel models with random intercepts by forum (no residual variance by topic was found in these models).

A5 Care and Finance Dictionaries

We utilize the dictionaries developed by Karpowitz and Mendelberg (2014) for care words and financial words. In their original dictionaries, two separate versions of the dictionaries were used and included many explicit variations of the tracked words for analysis in both LIWC and the TM package in R. Table A5 shows our variation on their dictionary. It includes nearly all of the words that they tracked, but has been modified to incorporate regular expressions that allow for simplification of the dictionaries and their implementation in almost any standard programming language without use of any specialized packages. For our analysis, these dictionaries were used in R using the gregexpr() syntax in base R.

Table A4 shows the results underlying Figure 4 in the main paper. For this table, the dependent variable is not logged or scaled.

Care Words
“\\<adolescent+”, “\\<babies”, “\\<baby”, “\\<boy+”, “\\<brother+”, “\\<care”, “\\<child+”, “\\<die+”, “\\<eat+”, “\\<education”, “\\<enough”, “\\<families”, “\\<family”, “\\<father+”, “\\<food”, “\\<girl+”, “\\<help+”, “\\<household+”, “\\<infant+”, “\\<juvenile+”, “\\<kid+”, “\\<kiddie”, “\\<kiddies”, “\\<kin”, “\\<live+”, “\\<minor+”, “\\<mother+”, “\\<mothers”, “\\<need+”, “\\<newborn+”, “\\<parent+”, “\\<poor”, “\\<poverty”, “\\<preteen+”, “\\<relative+”, “\\<safe”, “\\<safety”, “\\<school+”, “\\<sister+”, “\\<starv+”, “\\<struggl+”, “\\<student+”, “\\<surviv+”, “\\<teen+”, “\\<toddler+”, “\\<tot+”, “\\<welfare”, “\\<young+”, “\\<youth”)

Finance Words
“\\<affluent”, “\\<dollar+”, “\\<earn+”, “\\<elite+”, “\\<incentive+”, “\\<job+”, “\\<money+”, “\\<monied”, “\\<paid”, “\\<pay+”, “\\<rich”, “\\<salar+”, “\\<tax+”, “\\<wage+”, “\\<wealthy”, “\\<work+”

Dictionary based on a combination of Karpowitz and Mendelberg (2014) dictionaries for LIWC analysis and TM analysis. Additional characters are regular expression identifiers. “\\<” indicates that the word must start with the series of letters. “+” indicates that any set of further letters can end the word, allowing for plural forms and different tenses. Regular expression syntax is based on the base R gregexpr() syntax.

Table A4: The Effect of Participant Characteristics on the Use of Care and Financial Words

	Proportion Care Words	**Proportion Finance Words**
	Model 1	**Model 2**
Individual Level		
Female	0.007***	0.001
	(0.002)	(0.001)
31 to 45	0.002	0.00001
	(0.003)	(0.002)
46 to 64	−0.003	0.002
	(0.003)	(0.002)
65 or older	−0.006	−0.003
	(0.004)	(0.003)
Non-white	0.0005	−0.0002
	(0.002)	(0.001)
Forum Level		
N participants	0.002	−0.001
	(0.002)	(0.001)
Female moderator	0.002	0.001
	(0.003)	(0.003)
Unspecified moderator	−0.002	0.010***
	(0.003)	(0.003)
Proportion female	−0.001	−0.001
	(0.001)	(0.001)
Random Effect		
# of Forums	275	275
Forum Standard Deviation	0.009	0.011
# of Topics	20	20
Topic Standard Deviation	0.01	0.016
Constant	0.028***	0.017***
	(0.004)	(0.005)
N	1560	1560
Log Likelihood	3170.818	3508.229
AIC	−6315.636	−6990.458
BIC	−6246.054	−6920.876

***p < .01; **p < .05; *p < .1

Values in table are regression coefficients with standard errors in parentheses. The number of participants and proportion of female participants were scaled. All models are multilevel models with random intercepts by forum and topic.

A6 Methods Description and Additional Results: Content of Sessions and Chats

This section takes us through three different models for predicting the demographic characteristics of participants based on their word usage. For all of the models, we show the confusion matrices (Kennedy, 2015), which show the total numbers in each predicted category versus the actual values. The first models is a completely naive model, which serves as a baseline for understanding the performance of the other two models. In this model, word characteristics are not included as information, instead the modal category is always the predicted demographic category. If our more complex machine learning models return the same results, it indicates that there is no new information added by the word features.

The second model is a Naive Bayes model, which models the demographic category based on the independent probabilities of different word usages. One of the most intuitive applications of this is in Spam detection, where words like "viagra" or "watches" tend to indicate Spam is more likely. This simple classifier should be able to pick up on consistent and relatively large differences in the usage of particular words by different demographic groups. As is shown below, this does not provide a strong signal of demographic characteristics, and performs the same as the completely naive estimate.

Finally, we use support vector machines (SVM), which is a highly non-linear kernel-based estimator. While the features used to predict demographic categories are still individual words, SVM allows for highly non-linear and interactive patterns to emerge. As noted in the main paper, these are usually difficult to detect by humans reading text. In this case, the SVM model does outperform the completely naive estimator, but the differences in actual word usage for the words that are important for the model tend to be relatively small and the high non-linearity necessary for improvement on the completely naive model suggest that these differences would be difficult for a human to detect.

A6.1 Completely Naive Model

The completely naive models shows us the accuracy we would have from simply predicting the modal category for everyone in the session, and serves as a baseline for evaluation of subsequent models.

Table A5: Confusion Matrix for Completely Naive Model of Gender

	Reference	
Prediction	female	male
female	1073	592
male	0	0

Table A6: Confusion Matrix for Completely Naive Model of Age

	Reference	
Prediction	younger	older
younger	1053	565
older	0	0

Table A7: Confusion Matrix for Completely Naive Model of Race

	Reference	
Prediction	white	non-white
white	1100	882
non-white	0	0

A6.2 Naive Bayes

The naive Bayes classifier is a simple probabilistic classifier that is popular as a baseline model for text classification. In naive Bayes, the class label is given by the class C_k that maximizes the probability across N independent features (the frequency of words in the word matrix):

$$C_k = \underset{k \in \{male, female\}}{\arg\max} \; p(C_k) \Pi_{i=1}^{n} p(x_i | C_k) \tag{1}$$

The naive Bayes estimator is called naive because, as can be seen in equation (1), it makes a strong assumption about the independence of each feature. The likelihood is calculated independently for each feature (word) and is combined to produce the most likely classification. With this being said, naive Bayes has proven itself very efficient and useful in a variety of tasks, and is often difficult to outperform (Rennie et al., 2003).

As noted above, the naive Bayes classifier performs no better than our completely naive model.

Table A8: Confusion Matrix for Naive Bayes Model of Gender

	Reference	
Prediction	female	male
female	1073	592
male	0	0

Table A9: Confusion Matrix for Naive Bayes Model of Age

	Reference	
Prediction	younger	older
younger	1053	565
older	0	0

Table A10: Confusion Matrix for Naive Bayes Model of Race

	Reference	
Prediction	white	non-white
white	1100	882
non-white	0	0

A6.3 Support Vector Machines (SVM)

SVM is another commonly used method in text classification. SVM is an extension of the support vector classifier, which maximizes the margin between observations from different classes within a chosen level of error | this error tolerance becomes a primary tuning parameter for the model. Since only a few observations are relevant to establishing the maximum margin and the error, this method tends to be very efficient to compute and resistant to the influence of outliers. SVM extends the support vector classifier to a non-linear setting using kernels. Instead of being classified according to their values on particular features, classification is done through the observation's similarity to other observations | the kernel is a quantification of the similarity of two observations. A popular choice (and the one we use here) is the radial kernel, which takes the form:

$$K(x_i, x_{i'}) = exp(-\gamma \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2) \tag{2}$$

for each $j \in P$ features of the dataset. γ is a positive constant that can be used as a tuning parameter | as γ increases, the fit becomes more non-linear. Observations are then classified as a function $f(x_i)$ in the form:

$$f(x_i) = \beta_0 + \sum_{i' \in S} \alpha_{i'} K(x_i, x_{i'}) \tag{3}$$

Where the set S consists of the support vectors | the observations relevant to establishing the maximum margin and error | and $\alpha_{i'}$ is the weight of the observation.[1]

As with many flexible machine learning algorithm, SVM can produce any arbitrary decision boundary (i.e. we can produce perfect classification for any problem). Since such a complex decision boundary is unlikely to be generalizable and will make little substantive sense, SVM introduces a parameter that can be used for regularization | γ. Since a better fit on training data can always be obtained by adjusting γ, we utilize leave-one-out cross-validation to determine the error rate | for every $i \in N$ we run the model on the other $N - 1$ observations and use the results to predict the classification of i. The sum of the errors in predicting the withheld observation produces the model's cross-validation error.[2] The tuning parameters of the final model were were sigma = 0.0001908413 and tau = 0.0625.

In this case, the improvements to the completely naive model are significant (95% confidence intervals calculated by bootstrapping).

Table A11: Confusion Matrix for SVM Model of Gender

	Reference	
Prediction	female	male
female	1047	535
male	26	57

Table A12: Confusion Matrix for SVM Model of Age

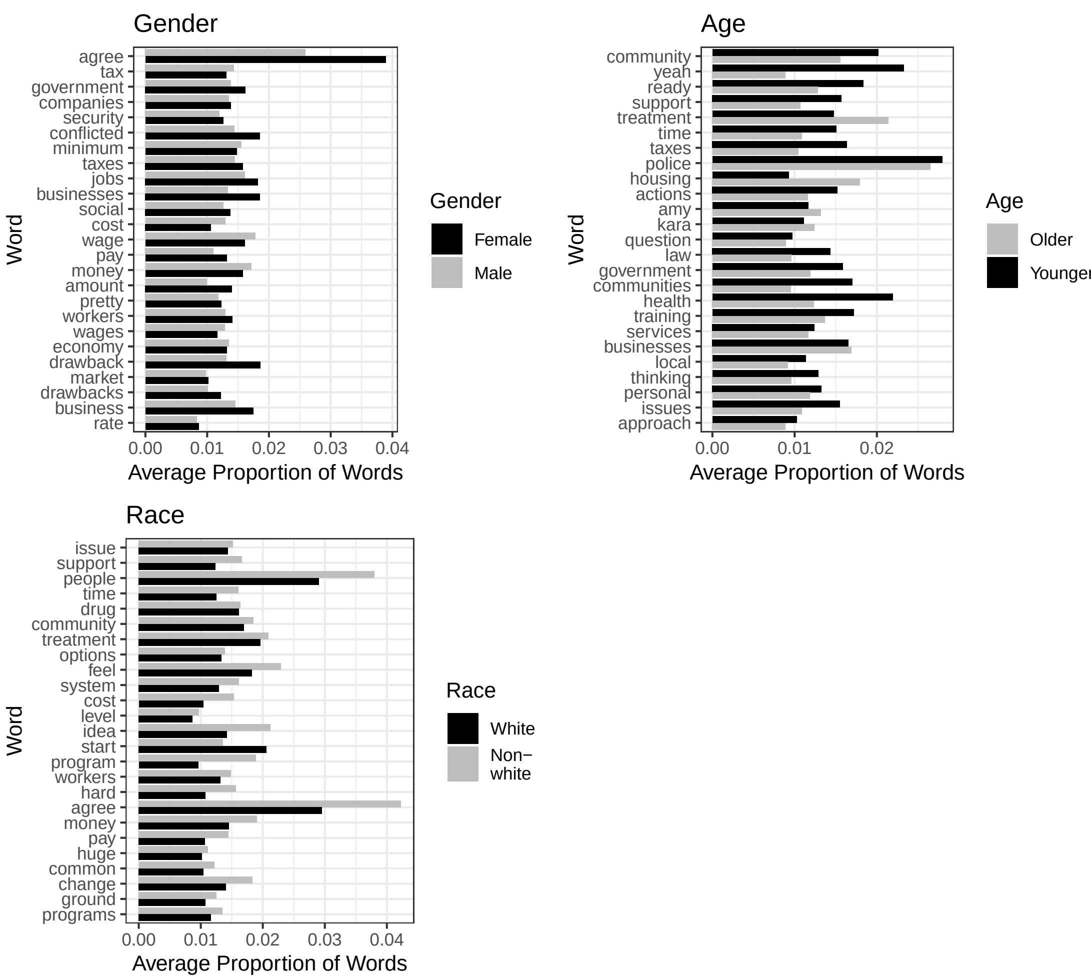
	Reference	
Prediction	younger	older
younger	976	350
older	77	215

Table A13: Confusion Matrix for SVM Model of Race

	Reference	
Prediction	white	non-white
white	859	363
non-white	241	519

Figure A4 shows the top-25 words in terms of their importance in the SVM model, with the most important words for distinguishing the categories plotted towards the top of the chart. The results presented herein are discussed in the main paper. The bars show the average proportion of total words are made up of these particular words for each category.[3]

Figure A4: Proportion of Words in Top-25 for Classification by Demographic Characteristics



Bar charts of top-25 most important features from SVM model of demographic groups on word frequencies. Bars are the proportion of non-stopwords made up of this word by demographic group.

A7 Effect of Female Majority

This section looks at whether there is a threshold effect in female participation. It is possible that our linear interactions in the main paper miss a discontinuity at a certain level of female participation, particularly when women come into a majority of participants in the session. As can be seen in Table A14, we find no such threshold effect when women become the majority of participants.

Table A14: Effect of Female Majorities in Session on Female Participation

	Chats	**Words**
	Model 1	**Model 2**
Individual Level		
Female	-0.168^{**}	-0.147^{*}
	(0.078)	(0.079)
31 to 45	0.320^{***}	0.273^{***}
	(0.081)	(0.081)
46 to 64	0.254^{***}	0.108
	(0.081)	(0.081)
65 or older	0.022	-0.153
	(0.122)	(0.121)
Non-white	-0.174^{***}	-0.303^{***}
	(0.049)	(0.049)
Forum Level		
N participants	-0.061	-0.024
	(0.058)	(0.053)
Female moderator	-0.103	-0.219^{**}
	(0.097)	(0.090)
Unspecified moderator	-0.756^{***}	-0.795^{***}
	(0.106)	(0.101)
Female majority	0.014	-0.028
	(0.099)	(0.094)
Cross-Level Interactions		
Female x female majority	0.074	0.107
	(0.098)	(0.099)
Random Effects		
# of Forums	275	275
Forum Standard Deviation	0.499	0.422
# of Topics	20	20
Topic Standard Deviation	0.105	0.15
Constant	0.742^{***}	0.862^{***}
	(0.106)	(0.107)
N	1609	1609
Log Likelihood	-2129.954	-2128.953
AIC	4287.907	4285.907
BIC	4363.275	4361.274

$^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

Values in table are regression coefficients with standard errors in parentheses. All models are multilevel models with random intercepts by forum and topic.

A8 Effect of Proportion of Non-white Participants and Over 65 Participants

Similar to "critical mass theory" for women, one might hypothesize that non-white and older participants become more active when their numbers in a session are greater. For our online CGA platform, this would seem to be unlikely, since race and age cannot be as accurately identified from a user's name and there are no associated visual cues to use in their place. For completeness, however, we tested whether there was an interaction effect between the level of participation for non-white and over-65 participants and the proportion they made up in a session. As Table A15 shows, we find no significant impact of either interaction.

Table A15: Testing Interaction with Proportion Non-white and Proportion over 65

	Chats	**Words**	**Chats**	**Words**
	Model 1	**Model 2**	**Model 3**	**Model 4**
Individual Level				
Female	-0.117^{**}	-0.075	-0.115^{**}	-0.074
	(0.046)	(0.047)	(0.046)	(0.047)
31 to 45	0.347^{***}	0.289^{***}	0.330^{***}	0.285^{***}
	(0.081)	(0.081)	(0.081)	(0.082)
46 to 64	0.281^{***}	0.122	0.267^{***}	0.124
	(0.082)	(0.081)	(0.083)	(0.082)
65 or older	0.051	-0.137	0.058	-0.099
	(0.122)	(0.122)	(0.135)	(0.137)
Non-white	-0.205^{***}	-0.325^{***}	-0.172^{***}	-0.303^{***}
	(0.052)	(0.052)	(0.049)	(0.049)
Forum Level				
N participants	-0.066	-0.026	-0.061	-0.026
	(0.058)	(0.053)	(0.058)	(0.053)
Female moderator	-0.099	-0.214^{**}	-0.090	-0.204^{**}
	(0.097)	(0.091)	(0.098)	(0.092)
Unspecified moderator	-0.811^{***}	-0.820^{***}	-0.744^{***}	-0.789^{***}
	(0.110)	(0.104)	(0.106)	(0.101)
Proportion non-white	0.118^{*}	0.054		
	(0.060)	(0.057)		
Proportion over 65			-0.028	-0.038
			(0.048)	(0.045)
Cross-Level Interactions				
Non-white x proportion non-white	-0.034	0.005		
	(0.074)	(0.073)		
Over 65 x proportion over 65			0.040	0.031
			(0.080)	(0.080)
Random Effects				
# of Forums	275	275	275	275
Forum Standard Deviation	0.495	0.42	0.499	0.421
# of Topics	20	20	20	20
Topic Standard Deviation	0.121	0.159	0.108	0.153
Constant	0.800^{***}	0.868^{***}	0.745^{***}	0.843^{***}
	(0.100)	(0.101)	(0.094)	(0.097)
N	1609	1609	1609	1609
Log Likelihood	-2129.085	-2129.692	-2130.974	-2130.020
AIC	4286.171	4287.385	4289.948	4288.040
BIC	4361.538	4362.752	4365.315	4363.407

$^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

Values in table are regression coefficients with standard errors in parentheses. All models are multilevel models with random intercepts by forum and topic.

A9 Disaggregation of Racial Categories

For some readers, the aggregation of racial categories into white/non-white may seem a little too reductive. In this section, we separate out the measured categories and report their separate values. As can been seen in Table A16, the main distinction does, indeed, appear to be between white and non-white participants. All minority groups participate at a lower rate in terms of both number of chats and total words in the sessions. Moreover, the substantive effects appear to be quite similar between groups, with only black and Latino participants showing a significant differentiation from each other, and only when it comes to the number of chats, with those identifying as black posting slightly more chats on average.

Table A16: Testing Disaggregation of Racial Categories

	Chats	Words
	Model 1	**Model 2**
Individual Level		
Female	−0.084***	−0.045*
	(0.028)	(0.025)
31 to 45	0.154***	0.073*
	(0.051)	(0.044)
46 to 64	0.128**	0.018
	(0.052)	(0.044)
65 or older	−0.170**	−0.272***
	(0.076)	(0.065)
Asian	−0.187***	−0.191***
	(0.048)	(0.041)
Latino	−0.214***	−0.186***
	(0.053)	(0.046)
Black	−0.149***	−0.164***
	(0.057)	(0.050)
Other/Mixed	−0.170***	−0.193***
	(0.044)	(0.038)
Forum Level		
N participants	−0.054	−0.033
	(0.036)	(0.029)
Female moderator	−0.045	−0.033
	(0.066)	(0.053)
Unspecified moderator	−0.544***	−0.319***
	(0.072)	(0.058)
Education	−0.129*	−0.081
	(0.070)	(0.057)
Stakeholder	−0.226	−0.129
	(0.174)	(0.164)
Survey	−0.033	0.044
	(0.093)	(0.079)
Random Effects		
# of Forums	275	275
Forum Standard Deviation	0.306	0.225
# of Topics	20	20
Topic Standard Deviation	0.069	0.104
Constant	0.945***	0.820***
	(0.087)	(0.076)
N	1609	1609
Log Likelihood	−1349.778	−1114.760
AIC	2735.557	2265.520
BIC	2832.457	2362.420

***p < .01; **p < .05; *p < .1

Values in table are regression coefficients with standard errors in parentheses. All models are multilevel models with random intercepts by forum and topic.

A10 Identification of Textual Gender Information

Although, as we note in the main paper, there is ample research to back up the idea that text signals of gender, provided through most CGA users' use of their first name as their name ID, some might still be skeptical about whether this is indeed the case. Previously, we attempted to test whether this was the case ourselves. We asked a group of 109 undergraduate students to read the transcripts from online town hall sessions with Members of Congress and answer a few factual questions to test their memory. In these sessions, similar to CGAs, the first name of the person asking questions was provided in the transcript. One of the factual questions we asked is if they thought the gender makeup of the sessions was skewed more towards men or women. *The students did little better than chance at identifying the gender makeup.*[4] Twenty-three of the students (about 21%) responded that they did not remember the relative distribution of men and women in the session, 40 (about 37%) guessed incorrectly, and 46 (about 42%) guessed correctly.

A11 Results Split by Educational Setting Sessions and Other Sessions

As we noted in the main text, this analysis combines data from sessions conducted with different purposes, and it is possible for there to be a relationship between the purpose and the relationships discussed in the main paper. We dealt with this in the main paper by including information on the type of session and including interactions with the type of session. In this section, we go a step further and evaluate separately the education CGA sessions and the non-education sessions. The results for the education sessions are shown in Table A17, while those for the non-education session are in Table A18.

For the effect of gender, the results remain largely the same – i.e., there appears to be little difference in participation between genders in both settings. The one difference observed is that the differences in the number of chats by women is not statistically significant in the non-education sessions. The findings on age are also a little different. He higher participation of those in the 31-45 age category is higher in both groupings, but is only statistically significant in the education setting. Conversely, the negative effect for participants in the 65 or older category primarily shows up in the non-educational setting, although this should not be too surprising, given that 3/4 of our over 65 respondents are in the non-educational setting. Finally, the finding on race is absolutely consistent across setting, with non-white participants speaking significantly less in both settings.

Table A17: The Effect of Participant Characteristics on the log Number of Chats and Words in CGA Sessions in Education Setting Only

	Chats	Words	Chats	Words	Chats	Words
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Individual Level						
Female	−0.109***	−0.046	−0.092**	−0.041	−0.138***	−0.057
	(0.040)	(0.036)	(0.042)	(0.038)	(0.048)	(0.043)
31 to 45	0.331***	0.174**	0.331***	0.174**	0.333***	0.175**
	(0.090)	(0.081)	(0.090)	(0.081)	(0.090)	(0.081)
46 to 64	0.183*	0.047	0.183*	0.046	0.185*	0.047
	(0.101)	(0.091)	(0.101)	(0.091)	(0.101)	(0.091)
65 or older	0.272*	0.076	0.282*	0.079	0.266*	0.074
	(0.144)	(0.130)	(0.144)	(0.130)	(0.144)	(0.130)
Non-white	−0.202***	−0.216***	−0.199***	−0.215***	−0.204***	−0.217***
	(0.038)	(0.035)	(0.038)	(0.035)	(0.038)	(0.035)
Preference deviation	0.126	0.013	0.130	0.014	0.130	0.014
	(0.162)	(0.146)	(0.162)	(0.146)	(0.162)	(0.146)
Forum Level						
N participants	−0.182***	−0.109**	−0.181***	−0.109**	−0.184***	−0.110**
	(0.050)	(0.043)	(0.050)	(0.043)	(0.050)	(0.043)
Female moderator	−0.148	−0.108	−0.150	−0.109	−0.201*	−0.129
	(0.091)	(0.089)	(0.092)	(0.089)	(0.103)	(0.098)
Unspecified moderator	−0.643***	−0.367***	−0.646***	−0.367***	−0.637***	−0.366***
	(0.090)	(0.083)	(0.090)	(0.083)	(0.089)	(0.083)
Proportion female	−0.088**	−0.085**	−0.133**	−0.098**	−0.089**	−0.085**
	(0.041)	(0.036)	(0.055)	(0.048)	(0.041)	(0.036)
Cross-Level Interactions						
Female x proportion female			0.070	0.021		
			(0.055)	(0.049)		
Female x female moderator					0.091	0.034
					(0.081)	(0.073)
Random Effects						
# of Forums	156	156	156	156	156	156
Forum Standard Deviation	0.321	0.253	0.324	0.254	0.32	0.253
# of Topics	16	16	16	16	16	16
Topic Standard Deviation	0	0.113	0	0.114	0	0.112
Constant	0.905***	0.802***	0.879***	0.794***	0.921***	0.809***
	(0.115)	(0.118)	(0.117)	(0.120)	(0.115)	(0.119)
N	898	898	898	898	898	898
Log Likelihood	−744.833	−648.413	−746.001	−650.420	−745.803	−650.004
AIC	1517.666	1324.825	1522.002	1330.839	1521.607	1330.009
BIC	1584.868	1392.027	1594.004	1402.842	1593.609	1402.011

***p < .01; **p < .05; *p < .1

Values in table are regression coefficients with standard errors in parentheses. Dependent variables have been scaled to aid in model convergence, so magnitude of coefficients is interpreted in standard deviations. The number of chats and number of words are also transformed by the natural log. The number of participants and proportion of female participants were also scaled. All models are multilevel models with random intercepts by forum and topic.

Table A18: The Effect of Participant Characteristics on the log Number of Chats and Words in CGA Sessions in Non-Education Setting Only

	Chats	**Words**	**Chats**	**Words**	**Chats**	**Words**
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Individual Level						
Female	−0.038	−0.022	−0.043	−0.023	−0.108*	−0.078
	(0.043)	(0.037)	(0.044)	(0.037)	(0.057)	(0.048)
31 to 45	0.064	0.016	0.065	0.016	0.053	0.007
	(0.060)	(0.048)	(0.060)	(0.048)	(0.060)	(0.048)
46 to 64	0.063	−0.018	0.062	−0.018	0.056	−0.022
	(0.058)	(0.046)	(0.058)	(0.046)	(0.058)	(0.046)
65 or older	−0.339***	−0.410***	−0.338***	−0.410***	−0.349***	−0.417***
	(0.087)	(0.070)	(0.087)	(0.070)	(0.087)	(0.070)
Non-white	−0.148***	−0.147***	−0.149***	−0.147***	−0.148***	−0.146***
	(0.047)	(0.038)	(0.047)	(0.039)	(0.047)	(0.038)
Preference deviation	−0.176	0.013	−0.170	0.014	−0.163	0.022
	(0.149)	(0.120)	(0.149)	(0.120)	(0.149)	(0.120)
Forum Level						
N participants	0.013	−0.016	0.012	−0.016	0.010	−0.018
	(0.049)	(0.034)	(0.049)	(0.034)	(0.049)	(0.034)
Female moderator	−0.057	−0.039	−0.058	−0.039	−0.162*	−0.122*
	(0.066)	(0.046)	(0.066)	(0.046)	(0.085)	(0.065)
Unspecified moderator	−0.150	−0.079	−0.152	−0.080	−0.145	−0.075
	(0.126)	(0.089)	(0.126)	(0.090)	(0.125)	(0.089)
Proportion female	0.065	0.037	0.098	0.047	0.068	0.039
	(0.048)	(0.034)	(0.069)	(0.053)	(0.048)	(0.034)
Cross-Level Interactions						
Female x proportion female			−0.049	−0.015		
			(0.074)	(0.061)		
Female x female moderator					0.160*	0.128*
					(0.084)	(0.070)
Random Effects						
# of Forums	119	119	119	119	119	119
Forum Standard Deviation	0.234	0.131	0.234	0.131	0.233	0.131
Constant	0.975***	0.841***	0.982***	0.843***	1.020***	0.877***
	(0.094)	(0.074)	(0.094)	(0.075)	(0.096)	(0.077)
N	711	711	711	711	711	711
Log Likelihood	−589.307	−450.395	−590.778	−452.250	−589.054	−450.470
AIC	1204.614	926.790	1209.556	932.501	1206.108	928.940
BIC	1263.980	986.157	1273.490	996.434	1270.041	992.874

***p < .01; **p < .05; *p < .1

Values in table are regression coefficients with standard errors in parentheses. Dependent variables have been scaled to aid in model convergence, so magnitude of coefficients is interpreted in standard deviations. The number of chats and number of words are also transformed by the natural log. The number of participants and proportion of female participants were also scaled. All models are multilevel models with random intercepts by forum. Topic random intercepts are omitted because of a lack of variance.

Supplementary Appendix: References

Karpowitz CF and Mendelberg T (2014) *The Silent Sex: Gender, Deliberation, and Institutions*. Princeton, NJ: Princeton University Press.

Kennedy R (2015) Making useful conflict predictions: Methods for addressing skewed classes and implementing cost-sensitive learning in the study of state failure. *Journal of Peace Research* 52(5): 649–664.

Rennie JD, Shih L, Teevan J and Karger DR (2003) Tackling the Poor Assumptions of Naive Bayes Text Classifiers. In: *Proceedings of the Twentieth International Conference on Machine Learning (ICML)*. Washington, DC: ICML. URL `http://people.csail.mit.edu/jrennie/papers/icml03-nb.pdf`.