

## WEB APPENDIX

### Speaking in Private: Privacy Expectations Depend on Communication Modality

Johann Melzner, Andrea Bonezzi, and Tom Meyvis

#### Overview

Web Appendix A – Studies 1-3 Additional Measures and Analyses	pp. 2–6
Web Appendix B – Google Search Study	pp. 7–9
Web Appendix References	p. 9

Survey materials, information on technical specifications, data, and syntax are available on OSF (<https://osf.io/4qjtz>) as well as in the Management Science replication package.

## 1. Study 1

### 1.1. Additional Measures

After participants answered the main dependent variable (i.e., the extent to which they found each of the five information flows acceptable), they also answered additional measures we administered for exploratory purposes that aimed to provide insight into potential underlying processes. In particular, they indicated the extent to which they felt like their answers made them identifiable (1=not at all; 9=very much), the extent to which they thought that the information they revealed about themselves in their answers was sensitive (1=not at all sensitive; 9=very sensitive), and the extent to which they felt comfortable answering the survey questions (1=not at all comfortable; 9=very comfortable).

We also extracted the word count for all responses using the lexical suite program (Rocklage and Fazio 2015) and calculated the mean word count for each participant. Note that, in this study, no audio content was collected, as we used a Java script integrated in Qualtrics that immediately transcribed spoken responses into text and recorded only the transcripts. As such, the word count for the speech and speech-transcription conditions was based on the automated transcripts

### 1.2. Results

*Identifiability.* A 3 (modality: text, speech, speech transcription) one-way ANOVA on perceived identifiability revealed a significant effect ( $F(2,172)=3.44, p=.034, \eta^2=.04$ ). Perceived identifiability in the speech condition ( $M=5.14, SD=2.51$ ) was marginally higher than in the text condition ( $M=4.42, SD=2.18, t(172)=1.68, p=.094, d=.31$ ) and significantly higher than in the speech transcription condition ( $M=4.04, SD=2.19, t(172)=2.58, p=.011, d=.48$ ). There was no statistically significant difference in perceived identifiability between the latter two conditions ( $t<1, p=.363, d=.17$ ).

*Perceived Sensitivity.* A 3 (modality: text, speech, speech-transcription) one-way ANOVA on perceived sensitivity revealed no significant effect ( $M_{Overall}=4.93, SD_{Overall}=2.08, F(2,172)=1.66, p=.193, \eta^2=.02$ ).

*Comfort.* Due to an error, the first 38 participants were not asked to indicate their level of comfort answering the survey questions, resulting in only 137 responses on this measure. A 3 (modality: text, speech, speech-transcription) one-way ANOVA on comfort revealed a significant effect ( $F(1,134)=6.38, p=.002, \eta^2=.09$ ). Comfort answering the questions was higher in the text condition ( $M=6.15, SD=2.46$ )

## SPEAKING IN PRIVATE

compared to both the speech condition ( $M=4.46$ ,  $SD=2.42$ ,  $t(134)=3.38$ ,  $p<.001$ ,  $d=.70$ ) and speech transcription condition ( $M=4.80$ ,  $SD=2.37$ ,  $t(134)=2.67$ ,  $p=.009$ ,  $d=.56$ ). The latter two conditions did not differ ( $t<1$ ,  $p=.507$ ,  $d=.14$ ).

*Word count.* A 3 (modality: text, speech, speech transcription) one-way ANOVA on word count revealed a significant effect ( $F(2,172)=4.38$ ,  $p=.014$ ,  $\eta^2=.05$ ). Word count in the speech transcription condition ( $M=4.16$ ,  $SD=2.12$ ) was significantly lower than in the text condition ( $M=6.50$ ,  $SD=5.16$ ,  $t(172)=2.93$ ,  $p=.004$ ,  $d=.54$ ) and marginally lower than in the speech condition ( $M=5.65$ ,  $SD=4.87$ ,  $t(172)=1.86$ ,  $p=.064$ ,  $d=.35$ ). There was no statistically significant difference in word count between the latter two conditions differ ( $t(172)=1.07$ ,  $p=.285$ ,  $d=.20$ ).

*Mediation analysis.* Out of these measures, only the pattern of perceived identifiability was consistent with the pattern of the dependent variable. Therefore, we conducted a mediation analysis using the process model 4 (Hayes 2017) with 10,000 bootstrap resamples. Modality served as a multi-categorical independent variable (speech=0; text=1, speech transcription=2), perceived identifiability as the mediator, and information flow acceptability (averaged across the 5 items;  $\alpha=.90$ ) as the dependent variable. No significant indirect effect of perceived identifiability emerged either for the speech and text ( $b=-.06$ ,  $SE=.07$ ,  $CI_{95\%}[-.23;.06]$ ) or for the speech and speech transcription comparison ( $b=-.10$ ,  $SE=.10$ ,  $CI_{95\%}[-.32;.08]$ ).

## 2. Study 2

### 2.1. Additional Measures

After participants answered the main dependent variable (i.e., the extent to which they found each of the five information flows acceptable), they also answered additional measures we administered for exploratory purposes that aimed to provide insight into potential underlying processes. In particular, participants indicated to what extent, while answering the survey questions, they expected that their responses would be transmitted somewhere else, saved and accessible at a later time, and accessible to other parties not present now (1=not at all; 9=very much). To the extent that participants perceived spoken responses to be more ephemeral than typed responses, they should find such information flows less likely. Thus, we averaged these measures into an ephemerality score. These measures loaded on a single factor and showed high internal consistency ( $\alpha=.83$ ). Participants also indicated the extent to which they feel like their responses made them identifiable (1=not at all identifiable; 9=very identifiable), the extent to which they thought that the information they revealed about themselves in their answers was sensitive (1=not at all sensitive; 9=very sensitive), the extent to which they felt comfortable answering the survey questions (1=not at all comfortable; 9=very comfortable), and the extent to which they thought the information they provided in their answers was valuable to others (1=not at all; 9=very).

We also extracted the word count for all responses using the lexical suite program (Rocklage and Fazio 2015) and calculated the mean word count for each participant. Note that, in this study, spoken responses were recorded using the service CameraTag integrated in Qualtrics. Semantic content contained in audio files<sup>1</sup> was automatically transcribed. Then research assistants blind to the study purpose and conditions checked the transcripts vis-à-vis the recordings and corrected them when necessary. Audio files were then removed to protect participants' privacy.

## 2.2. Results

*Ephemerality.* A 4 (modality: text, speech, voice-modification, recording-salience) one-way ANOVA on the ephemerality index revealed no significant effect ( $M_{\text{Overall}}=5.03$ ,  $SD=2.13$ ,  $F<1$ ).

*Perceived Identifiability.* A 4 (modality: text, speech, recording salience, voice modification,) one-way ANOVA on perceived identifiability revealed a significant main effect ( $F(3,219)=4.67$ ,  $p=.004$ ,  $\eta^2=.06$ ). Perceived identifiability was higher in the speech ( $M=5.29$ ,  $SD=2.21$ ) than in the text condition ( $M=3.98$ ,  $SD=1.91$ ,  $t(219)=3.25$ ,  $p=.001$ ,  $d=.61$ ). Perceived identifiability in the recording-salience condition ( $M=4.84$ ,  $SD=2.30$ ) did not differ from the speech condition ( $t(219)=1.11$ ,  $p=.270$ ,  $d=.21$ ) but was higher than in the text condition ( $t(219)=2.13$ ,  $p=.034$ ,  $d=.40$ ). Perceived identifiability in the voice modification condition ( $M=4.11$ ,  $SD=2.09$ ) did not differ from the text condition, ( $t<1$ ,  $p=.751$ ,  $d=.06$ ), but was significantly lower than in the speech ( $t(219)=2.89$ ,  $p=.004$ ,  $d=.55$ ), and marginally lower than in the recording salience condition ( $t(219)=1.79$ ,  $p=.075$ ,  $d=.34$ ).

*Perceived Sensitivity.* A 4 (modality: text, speech, recording salience, voice modification) one-way ANOVA on perceived sensitivity revealed no significant effect ( $M_{\text{Overall}}=4.42$ ,  $SD=2.18$ ;  $F<1$ ).

*Comfort.* A 4 (modality: text, speech, voice modification, recording salience) one-way ANOVA on comfort revealed no significant effect ( $M_{\text{Overall}}=5.98$ ,  $SD=2.16$ ;  $F<1$ ).

*Perceived information value.* A 4 (modality: text, speech, recording salience, voice modification) one-way ANOVA on perceived information value revealed no significant effect ( $M_{\text{Overall}}=4.84$ ,  $SD=2.03$ ;  $F(3,219)=1.50$ ,  $p=.217$ ,  $\eta^2=.02$ ).

*Word count.* A 4 (modality: text, speech, recording salience, voice modification) one-way ANOVA on word count revealed a significant effect ( $F(3,210)=13.48$ ,  $p<.001$ ,  $\eta^2=.16$ ). Word count was lowest in the text condition ( $M=2.98$ ,  $SD=1.59$ ), followed by the speech condition ( $M=5.18$ ,  $SD=3.75$ ,  $t(210)=3.14$ ,  $p=.002$ ,  $d=.60$ ), the recording salience condition ( $M=6.29$ ,  $SD=4.22$ ,  $t(210)=1.55$ ,  $p=.123$ ,  $d=.30$ ), and the voice modification condition ( $M=7.17$ ,  $SD=4.48$ ,  $t(210)=1.23$ ,  $p=.219$ ,  $d=.24$ ).

---

<sup>1</sup> Recordings from nine participants were unavailable. This was likely due to participants using a browser (version) that was incompatible with Cameratag, rather than the recommended browser (most up-to-date Chrome). Due to COVID 19, participants completed this study at home, rather than in private laboratory rooms.

## SPEAKING IN PRIVATE

*Mediation analysis.* Out of these measures, only the pattern of perceived identifiability was consistent with the pattern of the dependent variable. Therefore, we conducted a mediation analysis using the (process model 4; Hayes 2017) with 10,000 bootstrap resamples. Modality served as a multi-categorical independent variable (text=0; speech=1, voice modification= 2, and recording salience=3), perceived identifiability as the mediator, and information flow acceptability (averaged across the 5 items;  $\alpha=.89$ ) as the dependent variable. No significant indirect effect of perceived identifiability on information flow acceptability emerged neither for the text–speech ( $b=-.04$ ,  $SE=.11$ ,  $CI_{95\%}[-.28;.17]$ ), the text–voice modification ( $b=.004$ ,  $SE=.04$ ,  $CI_{95\%}[-.09;.07]$ ), nor the text–speech recording salience comparison ( $b=-.03$ ,  $SE=.08$ ,  $CI_{95\%}[-.21;.12]$ ).

### 2.3. Post-Test

Given that distortion and pitch-shift was used as manipulation in the voice-modification condition, a post-test assessed the intelligibility of the modified spoken content of the audio file: “I bought an Ipad. I really like it. I use it a lot.” To this end 180 participants from the same population used for study 2 were presented with the audio file and asked to type what they heard. Three participants provided nonsensical transcripts (e.g., “lolololol”) and another three reported problems hearing or did not answer the question. Therefore, 174 participants remained for analysis (gender: 87 female, 86 male, 1 did not want to specify; Age:  $M=19.63$ ,  $SD=1.15$ ).

For each participant, we counted the number of correctly transcribed words and divided it by the total number of words (=13) to calculate the percentage of correctly transcribed words. On average, participants transcribed 81.61% ( $\approx 11/13$  words) of the content correctly ( $SD=17.21\%$ ). 36.2% of participants transcribed the entire content (i.e., all 13 words) correctly. Overall, these results suggest that the content was intelligible, despite the pitch-shift and distortion applied.

## 3. Study 3

### 3.1. Additional Measures

After participants answered the main dependent variable (i.e., the extent to which they found each of the five information flows acceptable) and the measures of ownership and identifiability reported in the main manuscript, they also answered additional measures we administered for exploratory purposes. In particular, participants indicated to what extent, while answering the survey questions, they expected that their responses would leave a permanent record, be accessed by somebody else who is not present now, and be transmitted somewhere (1=not at all; 9=very much). We averaged these measures into an ephemerality score. These measures indeed loaded on a single factor and showed adequate internal consistency ( $\alpha=.70$ ). Participants also indicated the extent to which they thought that the information they revealed about themselves in their answers was sensitive (1=not at all sensitive; 9=very sensitive), the

## SPEAKING IN PRIVATE

extent to which they felt comfortable answering the survey questions (1=not at all comfortable; 9=very comfortable), how much information about themselves they thought they provided in their responses (1=not much information at all; 9=a lot of information), and the extent to which they thought that the way in which they typed or spoke their responses revealed something about themselves (1=not at all; 9=very much).

We also extracted the word count for all responses using the lexical suite program (Rocklage and Fazio 2015) and calculated the mean word count for each participant. Note that, in this study, spoken responses were recorded using the service CameraTag integrated in Qualtrics. Semantic content contained in audio files were automatically transcribed. Then research assistants blind to the study purpose and conditions checked the transcripts vis-à-vis recordings and corrected them when necessary. Audio files were then removed to protect participants' privacy.

### 3.2. Results

*Ephemerality.* A 2 (modality: text, speech) one-way ANOVA on the ephemerality index revealed no significant effect ( $M_{\text{Overall}}=5.83$ ,  $SD=1.91$ ,  $F<1$ ).

*Perceived Sensitivity.* A 2 (modality: text, speech) one-way ANOVA on perceived sensitivity revealed no significant effect ( $M_{\text{Overall}}=4.16$ ,  $SD = 2.27$ ,  $F(1,111)=1.14$ ,  $p=.287$ ,  $\eta^2=.01$ ).

*Comfort.* A 2 (modality: text, speech) one-way ANOVA on comfort revealed a significant effect ( $F(1,111)=5.76$ ,  $p=.018$ ,  $\eta^2=.05$ ). Comfort answering the survey questions was higher in the text ( $M=6.54$ ,  $SD=2.12$ ) versus speech condition ( $M=5.51$ ,  $SD=2.42$ ).

*Perceived information quantity.* A 2 (modality: text, speech) one-way ANOVA on perceived information quantity revealed no significant effect ( $M_{\text{Overall}}=4.51$ ,  $SD=1.94$ ;  $F<1$ ).

*Perceived paralinguistic information.* A 2 (modality: text, speech) one-way ANOVA on perceived paralinguistic information revealed a significant effect ( $F(1,111)=4.32$ ,  $p=.040$ ,  $\eta^2=.04$ ). Participants in the speech condition ( $M=6.02$ ,  $SD=2.16$ ) thought that the way they spoke responses revealed more information about them than participants in the text condition thought that the way they typed responses did ( $M=5.20$ ,  $SD=2.04$ ).

*Word count.* A 2 (modality: text, speech) one-way ANOVA on average word count revealed no significant effect ( $M_{\text{Overall}}=5.85$ ,  $SD=3.82$ ;  $F<1$ ).

*Mediation analysis.* Because there were significant differences in comfort and perceived paralinguistic information as a function of modality, we conducted another mediation analysis including these measures as mediators in addition to ownership and identifiability. Specifically, we conducted a mediation analysis using the Process Macro Model 4 (Hayes 2017) with 10,000 bootstrap resamples. Modality served as the independent variable (text=0; speech=1), information flow acceptability (averaged across the five items;  $\alpha=.87$ ) as the dependent variable, and ownership, identifiability, comfort, and

## SPEAKING IN PRIVATE

perceived paralinguistic information as mediators. This analysis revealed a significant indirect effect of ownership ( $b=-.41$ ,  $SE=.20$ ,  $CI_{95\%}[-.86;-.09]$ ), but no significant indirect effects of identifiability ( $b=.13$ ,  $SE=.15$ ,  $CI_{95\%}[-.14;.45]$ ), comfort ( $b=-.14$ ,  $SE=.11$ ,  $CI_{95\%}[-.41;.02]$ ), or perceived paralinguistic information ( $b=.11$ ,  $SE=.10$ ,  $CI_{95\%}[-.02;.35]$ ). When controlling for the indirect effects, the direct effect of modality remains significant ( $b=-.80$ ,  $SE=.36$ ,  $CI_{95\%}[-1.51;-.09]$ ,  $p=.027$ ), indicating partial mediation.

## WEB APPENDIX B – Google Search Study

Survey materials, data, and syntax are available on OSF (<https://osf.io/4qjgz>) as well as in the Management Science replication package.

### 1. Study Goal

This study aimed to test a practically relevant consequence of our basic finding that consumers find information flows less acceptable when information is disclosed via speech as compared to text. Consumers are commonly targeted with ads based on information they provide in a different context. That is, targeted ads are often displayed by a different entity than the one the information was disclosed to, presented sometime after the information has been disclosed, and presented on a different device and/ or website than the one information was disclosed on. As demonstrated in studies 1–3, consumers find such information flows less acceptable for information disclosed through speech versus text. Thus, we reasoned that consumers should feel that a targeted ad violates their privacy more when it is based on information they disclosed in a different context via speech versus text. In turn, such increased feelings of privacy violations may foster lower receptivity towards the targeted ad. We tested this prediction in a practically relevant context: Google searches.

### 2. Procedure

Participants were individually seated in private laboratory rooms and completed this study as part of an one hour experimental session. In part 1 of the study, participants were told to imagine that they wanted to put together an itinerary for a trip to Atlanta. For this purpose, participants were instructed to search for five things to do in Atlanta using Google. Participants were assigned to one of two conditions: text-search and speech-search. In the text-search condition, participants were instructed to enter search terms in Google by typing. In the speech-search condition, participants were instructed to enter search terms by speaking into the built-in laptop-microphone using the Google voice functionality. In the speech-search condition the keyboard was disabled to ensure that participants used only the voice functionality.

## SPEAKING IN PRIVATE

Participants were asked to look for a 1. sight, 2. museum, 3. restaurant, 4. bar or club, and 5. show (comedy, concert, musical, or play), and write down their selection on a piece of paper placed next to the laptop. Participants were given five minutes for this task (a timer was displayed on the survey page for them to check the time).

Afterwards, participants were given a randomly generated three-digit code (numbers from 101-200 [201-300] were generated for the speech-search [text-search] condition) that they were instructed to write down on a post-it next to them. Participants were instructed to get the research assistant's attention. The research assistant entered participants' private rooms, opened a second laptop and instructed participants to complete the rest of the studies in the lab session using that second laptop. Participants completed about 45 minutes of unrelated studies before starting part 2 of this study.

At the beginning of part 2, participants were instructed to enter the three-digit code provided at the end of part 1. After entering the code, participants were informed that the search terms they had used for their Google searches in the first study of the lab session (i.e., part 1 of this study) were saved and that an ad was now being generated based on inferences drawn from those search terms. After 20 seconds, participants were presented with an ad for a restaurant called Eatery 21 (adopted from Summers, Smith, and Reczek 2016). Presenting an ad in the Qualtrics environment based on search terms disclosed to Google, after a 45-minute time delay, and on a different device than the one they used for the Google searches, was meant to make information flows not prescribed by the initial disclosure context salient.

Following the presentation of the ad, participants indicated the degree to which they feel like their privacy has been violated (1=not at all; 7=very much). As a measure of ad receptivity, participants stated their agreement with the statements "I am interested in visiting the website of Eatery 21" and "I am interested in dining at Eatery 21" (-3=strongly disagree; 0=neither agree nor disagree; +3= strongly agree; coded from 1-7). These two items were highly correlated ( $r=.83$ ) and averaged into an index of ad receptivity. In order to ensure that differences in ad receptivity were not a function of differences in how well the ad matched the search terms participants used in their Google searches for part 1 of the study, participants indicated how good a match the personalized ad was for them based on their Google search terms (1=not at all a good match; 7=a very good match).

174 participants completed both parts 1 and 2 (Gender: 87 female, 86 male, 1 did not indicate; Age:  $M=20.06$ ,  $SD=1.51$ ).

### 3. Results

*Privacy violation.* A 2 (modality: Google text search, Google speech search) one-way ANOVA on privacy violation revealed a marginally significant effect ( $F(1,172)=3.74$ ,  $p=.055$ ,  $\eta^2=.02$ ). Participants indicated higher privacy violations in the Google speech search ( $M = 4.08$ ,  $SD = 1.97$ ) versus the Google text search condition ( $M=3.52$ ,  $SD=1.84$ ).

## SPEAKING IN PRIVATE

*Ad receptivity.* A 2 (modality: Google text search, Google speech search) one-way ANOVA on ad receptivity revealed a significant effect ( $F(1,172)=4.55, p=.034, \eta^2=.03$ ). Participants indicated reduced interest in the restaurant displayed in the ad in the speech-search ( $M=3.47, SD=1.44$ ) versus text-search condition ( $M=3.94, SD=1.51$ ). Including ad fit as covariate renders the effect marginally significant ( $F(1,171)=3.84, p=.052, \eta^2=.02$ )

*Ad fit.* A 2 (modality: Google text search, Google speech search) one-way ANOVA on ad fit revealed no significant effect ( $M_{\text{Overall}}=3.13, SD=1.64, F<1$ ).

*Mediation analysis.* We conducted a mediation analysis using the Process Macro Model 4 (Hayes 2017) with 10,000 bootstrap resamples. Modality served as independent variable (Google text search=0; Google speech search=1), privacy violations as mediator, ad fit as a covariate, and ad receptivity as dependent variable. This analysis revealed a significant indirect effect of privacy violation ( $b=-.09, SE=.05, CI_{95}[-.210;-.003]$ ), rendering the direct effect insignificant ( $b=-.29, SE=.19, CI_{95}[-.66;.09], p=.133$ ).

#### 4. Discussion

The results of this study show a practically relevant consequence of our basic finding that consumers find information flows less acceptable when information is disclosed via speech as compared to text. Participants felt that an ad targeted based on information flows not prescribed by the disclosure context violated their privacy more when the information was disclosed via speech versus text. In turn, increased feelings of privacy violations lead to lower receptivity towards the targeted ad.

## WEB APPENDIX – References

- Hayes AF (2017) *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, Guilford publications.
- Rocklage MD, Fazio RH (2015) The lexical suite: Adjective use as a means of assessing and distinguishing attitude valence, extremity, and emotionality. *J Experimental Social Psy.* 56:214–227.
- Summers CA, Smith RW, Reczek RW (2016) An audience of one: Behaviorally targeted ads as implied social labels. *J. Consumer Res.* 43(1):156–78.