Interlocutor relation predicts the formality of the conversation

An experiment in American and British English

Stephanie Rotter[a] & Mingya Liu[a]

[a]Humboldt-Universität zu Berlin

Formality is an important notion in register research. However, it is hard to give it a precise definition. Furthermore, it is challenging to identify the situational parameters required for levels of formality. Potential parameters include the social relation between interlocutors. We conducted two rating studies with American and British English speakers and tested the connection between formality and social relations, while exploring the influence of certain demographic factors on the rating. We hypothesised that public social relations are associated with more formal contexts than private social relations. The results strongly confirmed the hypothesis. We expect that our findings shed further light on the connection between formality and social relations and consequently, inform experimental set-ups investigating variation in language use.

Keywords: Formality, situational parameter, interlocutor relation, rating study, American and British English

1 Introduction

Formality is an important notion in register research. Certain situations, e. g. a court hearing, are typical examples of formal settings, while others, e. g. a private supper with family and friends, are perceived as rather informal settings in various cultures (Irvine 1979, Heylighen & Dewaele 1999). While it is comparatively straightforward to recognise markers of formal or informal speech, i. e. the usage of standard vs. non-standard varieties of a language (Ferguson 1959, Creber & Giles 1983), it is a challenging task to identify the extra-linguistic aspects or parameters which distinguish formal from informal settings. This paper aims at contributing to this issue with two rating studies investigating the influence

of social relations and the perceived level of formality in American and British English.

Formality describes a degree of attention paid to the rules of social conduct (Irvine 1979, Labov 1972, Pescuma et al. 2023), which includes linguistic discourse and any behaviors in social interaction. Thereby 'formal' denotes high degrees of socially shared structure rendering the interaction predictable, while the opposite 'informal' refers to lower degrees of structures and rules. High degrees of formality can be found in religious rituals (Irvine 1979) and lower degrees in casual events like shopping (Ferguson 1959). Formality is associated with politeness and appropriateness whereas informality promotes intimacy and emotionality among the interlocutors (Brown 2015, Heylighen & Dewaele 1999, Irvine 1979, Halliday 1964). Formality acts on two levels which interact with one another, namely extra-linguistic and intra-linguistic aspects. Extra-linguistic aspects can suggest the appropriate formality degree of the upcoming speech, e. g. formal speech can be suggested in a court hearing by the rule-governed procedure and people wearing suits among others. Formal markers in the speech have an influence on the perceived level of the situation, e. g. when meeting a stranger who approaches us with a large amount of formal markers in the speech. This paper focuses on the first aspect, namely how language comprehenders evaluate situations and anticipate the level of formality in the upcoming speech.

The observation that linguistic behavior varies with respect to the degree of expected rules contributed to the study of language variation and registers. In traditional variationist approaches which focus on situational aspects, formality represented the key dimension to differentiate contextual styles (or register) (Labov 1972). Other functional approaches included aspects of formality in their dimensions, e. g. in "Dimension 1 involved vs. informational production" (Heylighen & Dewaele 1999, Biber & Finegan 1994, Biber 1988). This dimension contrasts linguistic features rendering the discourse content interactional (e. g. in a face to face conversation) or dense in information (e. g. in written academic texts) (Biber 2009).

We consider language register with a combined definition as the "socially recurring intra-individual variation that are influenced by situational and functional settings" (Lüdeling et al. 2022: 3). Formality still remains a crucial concept as one dimension of the situational-functional settings and is hence vital for a deeper understanding of registers. We focus on the situational aspects of formality in the remainder of the paper.

As a dimension, formality is a product of various situational settings or parameters and their interaction, among others, the modality of the text or speech, the setting of the conversation, and social aspects of interlocutors (Agha 2007, Biber

& Finegan 1994). The manipulation of one of such parameters while keeping the others relatively stable, sometimes suffices to suggest different degrees of formality shown by more frequent use of linguistic formality markers formal and less frequent use for informal situations. Corpus studies found that the formality differed across text modality: written texts show a higher level of formality than spontaneous speech marked by more complex grammatical features (Biber et al. 2007, Halliday 1964). In a rating study the concrete location of interactions was manipulated. The results revealed that the school is perceived as a more formal setting, requiring rather formal speech, in comparison to a youth club (Creber & Giles 1983). Furthermore, social aspects of the interlocutors were shown to impact the formality of speech: The social distance to the interlocutor changes the frequency of formality markers in speech (e. g. of laughter and interjections) in that more markers were used when the social distance to the interlocutor was large, i. e. interactions with an unfamiliar person, in comparison to a smaller distance, i. e. interactions with a familiar person (Koppen et al. 2019). Besides distance, other social aspects affect the level of formality as well, e. g. social hierarchy or status of interlocutors (Agha 1998).

The explicit configuration of parameters is culturally conventionalized (Berruto 2010, Agha 2007, Paolillo 2000, Atkinson 1982, Irvine 1979). This close connection to culture has the effect that geographically distinct communities of the same language can have different versions of formal registers (see Berruto 2010 for a detailed discussion) and also differing parameter requirements for formal settings (Agha 2007, Irvine 1979). Some varieties of English, e. g. American and British English, developed in different sociohistorical contexts, thus, the understanding of formality can differ among English speakers.

We were interested in the effect of the social parameters on the perceived level of formality in two communities of the same language, i. e. American and British English speakers. More specifically, we ask how language comprehenders evaluate the situation and anticipate the language markers. Furthermore, we were interested in whether the manipulation of the social relations, i. e. public vs. private, suffices to evoke a formal or informal context in rating studies. We hypothesised that American and British English speakers would rate the formality degrees higher in public social relations, e. g. with the *boss*, than in private social relations, e. g. with the *mother*. For exploratory reasons, we additionally analysed demographic factors, i. e. age, gender, environment people grew up in, as well as education level, assuming that rating variation can be partially explained by these factors.

In order to test the hypotheses, we conducted two rating studies in which native speakers of American and British English provided their estimation about

the formality level of the next interaction in a controlled situation. The results across both data sets showed strongly that public social relations were associated with more formal contexts than private social relations.

Having summarized relevant literature, we present the experimental approach we used for both experiments and their results in Section 2. The results are discussed in Section 3. Section 4 concludes the paper.

2 Experimental Method

The aforementioned hypotheses were tested with two rating studies with American and British English participants. In Section 2.1 and 2.2, we report on the experimental set-up and the results of the American experiment. In Section 2.3, we focus on those aspects of the experimental set-up of the British experiment which differed from the American experiment and we report on the results in Section 2.4.

2.1 Experiment 1: American English

We first describe the design and material, as well as the procedure. Then, we report on the demographic background of the participants. Lastly, we describe the statistical approach used for the data analysis.

2.1.1 Design and Material

The experiment was implemented using PCIbex (Zehr & Schwarz 2018) and hosted on the *PCIbex farm* platform (https://farm.pcibex.net/). The experiment used a one factorial design with the factor SOCIAL RELATION (public vs. private). We used 48 critical items (see example 1) and 16 fillers of similar structure. The items consisted of short stories with the following structure: (S1) introduced the agent of the story in a general context. (S2) encoded the formality manipulation via social relations. For the public condition, we used 24 labels referring to social relations in a working environment, e. g. *boss* or *costumer*. For the private condition, we used 24 labels referring to intimate relationships in the private sector, e. g. *wife* or *friend*.

Each label was used twice across the items such that the agent in (S1) was once a female and once a male agent as indicated by a proper name. (S1) and (S2) were

followed by (Q1), asking whether the introduced agent is going to talk formally[1] to the indicated person. (Q2) represented a comprehension question. In half of the items, participants additionally answered a comprehension question.

(1) (S1) George Henderson works in a shop.

     (S2) He says to his **boss**$_{\text{PUBLIC}}$: …

           He says to his **sibling**$_{\text{PRIVATE}}$: …

     (Q1) Is George Henderson going to talk formally?

     (Q2) Does George Henderson work in a shop?

2.1.2 Procedure

Participants first gave their informed consent before accessing the instructions to the experiment. After the participant read the introduction, three practice stories followed in order to familiarize the participants with the set-up. Then, the experimental part started. Participants indicated that they had read the sentence by pressing on the space key before the next sentence appeared in the middle of the screen. (S1) and (S2) were shown separately one after the other. (S2) remained on the screen when (Q1) together with the scale appeared. The answer to the question was given on a 7-point Likert scale in which all points were labelled, i. e. "Certainly not/yes" (1/7), "Probably not/yes" (2/6), "Possibly not/yes" (3/5), "unclear" (4). Participants used the according keys to indicate their answer. (Q1) was replaced by (Q2) – a binary comprehension question with the options (1) "No" and (7) "Yes". Again, participants used the according key to indicate their answer. After each story, a fixation cross appeared in the middle of the screen for 400ms in order to facilitate a proper distinction between the stories. After the experimental part, the participants took part in a short survey about their demographic and language background.

---

[1] The formality measure can be seen as similar to the acceptability measures commonly used in linguistics (e. g. for grammaticality, acceptability, naturalness and appropriateness judgements) in that they all collect language users' intuitions about language and language use. In the present paper, we used the formality ratings as a direct measure of language users' perception of conversational contexts; in this regard, they pattern more similarly to the measures of social meanings about the speaker, e. g. whether they are being friendly or polite via what they say and how they say it. Future work is needed to explore the usefulness and the scope of the formality measure in register research

2.1.3 Participants

100 American English speakers (female: 47, male: 49, non-binary: 4) were recruited through the crowd-sourcing platform Prolific (https://www.prolific.co/) and received monetary compensation for their participation. The experiment took roughly 15 minutes. We received informed consent from all participants.

We removed two participants who acquired English in school. All participants matched the inclusion age range. Based on the results of the comprehension question, we removed the data of five participants with an accuracy below 90 %.

The remaining data consisted of 93 participants with the mean age of 37.9 years (SD=11.1) ranging from 19 to 64 years. Eight participants reported to be dialect speakers (i. e. African American Vernacular English as well as Midwest, Southeastern, Southern, and Western Pacific American English dialect). More than the majority (60 %) of the participants graduated from college. One quarter (26 %) obtained a high school degree, 13 % have a graduate degree and 1 % do not have a high school degree. More than half of the participants (59 %) were raised in a suburban environment, while 25 % grew up in a rural and 16 % in an urban environment.

2.1.4 Statistics

The data was processed and analyzed with the software R (R Core Team 2022) in the RStudio envionment (RStudio Team 2022). We removed the data from participants who did not match the inclusion criteria of being English native speakers or aged between 18 to 65 years. For the outlier removal, we used the answers of the comprehension question (Q2) and the reading times of (S1) and (S2). We removed the data from participants with an accuracy rate below 90 % of the 24 comprehension questions. We removed entire trails if the reading time of (S1) or (S2) was below 400ms.

The remaining data was plotted using the package 'gglot2' (Wickham 2016). We first report on the overall distribution of the mean and median ratings per condition. Furthermore, we report on the descriptive statistics of specific labels with high or low extreme values in comparison to the majority of the other labels.

The data was analysed using the package 'ordinal' (Christensen 2019) in the cumulative link function model framework[2]. The link functions were identified by comparing the loglikelihood values of each of the five possible link functions (i. e. probit, logit, cauchit, loglog, and cloglog) with each other (Christensen 2011).

---

[2]See Howcroft & Rieser (2021) and Liddell & Kruschke (2018) for further discussions on using ordered regressions for ordinal data.

This method has the advantage that the chosen link function follows the shape of the data and accounts for skewness which maximizes the model's fit. The logit function had the highest value and was thus used for the models.

Random effect structures were obtained with the help of the most parsimonious model approach (Bates et al. 2015); the used models are indicated in the result section. For the main analysis, we sum coded the following predictors: SOCIAL RELATION[3] with private (0.5) and public (-0.5), as well as demographic information from participants, i. e. GENDER[4] with female (0.5) and male (-0.5), EDUCATION and the ENVIRONMENT participants grew up in. The latter two had three levels and were coded the following way: EDUCATION[5] with college (c, 0.5) and high school/graduate (hs/g, -0.5), as well as ENVIRONMENT with the levels suburban (su, 0.5) and rural/urban (r/u, -0.5). Additionally, we added AGE as a continuous predictor and the 2-way interactions of SOCIAL RELATION with the six demographic predictors. If interactions turned out significant, we conducted sub-analyses with the data of the condition public and private. We used the following predictors: AGE, GENDER, EDUCATION, and ENVIRONMENT; the coding remained identical as previously stated. We only report on effects of the sub-analyses which showed an interaction effect in the main analyses.

P-values in all analyses were obtained with the help of loglikelihood ratio test comparisons of nested models. All statistical values of means, estimates and there like, are rounded to the second decimals except for p-values smaller than 0.01.

2.2 Results

190 entire trails were removed from the data set based on the reading times. The descriptive statistics of the remaining data across the conditions are shown in Table 1. Figure 1 depicts the means, medians, and standard error of each label. Means and medians of the public condition range from 5 to 6. Certain labels, i. e. *chef* (mean = 4.63, SD = 1.46, median = 5), *salesperson* (mean = 4.42, SD = 1.35, median = 5), and *waiter* (mean = 4.87, SD = 1.38, median = 5) received the lowest means and medians. *General* (mean = 6.81, SD = 0.45, median = 7) and *judge* (mean = 6.71, SD = 0.72, median = 7) received the highest mean and median.

---

[3]Factors are named with capital letters.

[4]The factor GENDER initially had three levels, i. e. female, male, and non-binary. However, given the sparse data for the level "non-binary", we excluded it from the analysis by coding it with 0.

[5]The factor EDUCATION initially had four levels, i. e. graduate, college, high school, and no high school degree. However, given the sparse data for the level "no high school", we excluded it from the analysis by coding it with 0.

In the private condition, most means and medians range from 1.8 to 2.5 with the exception of the label *confidant* (mean = 3.07, SD = 1.75, median = 3) and *father* (mean = 2.82, sd = 1.45, median = 2).

Table 1: Descriptive statistics of the ratings of the American data set.

Social relation	N	Median	Mean	SD	SE
public	2151	6	5.57	1.26	0.03
private	2123	2	2.19	1.24	0.03

For the main analysis, the model with random item intercepts fit best the data. The output of the model is depicted in Table 2. The results showed a significant main effect of SOCIAL RELATION in that private conditions received lower formality ratings than public conditions ($\hat{\beta}$ = -5.12, $\chi^2(1)$ = 521.66, p < 0.0001). In addition, AGE showed a significant main effect in that the higher the age, the lower the rating ($\hat{\beta}$ = -0.01, $\chi^2(1)$ = 7.56, p = 0.006). The main effect GENDER turned out significant in that females provided lower ratings than males ($\hat{\beta}$ = -0.29, $\chi^2(1)$ = 24.78, p < 0.0001). ENVIRONMENT$_{(su\ vs.\ u)}$ showed a significant main effect in that people who grew up in a suburban environment provided lower ratings than people who grew up in an urban environment ($\hat{\beta}$ = -0.53, $\chi^2(1)$ = 24.73, p < 0.0001). Furthermore, the following 2-way interaction turned out significant: SOCIAL RELATION and AGE ($\hat{\beta}$ = 0.02, $\chi^2(1)$ = 9.26, p = 0.002), SOCIAL RELATION and GENDER ($\hat{\beta}$ = -0.50, $\chi^2(1)$ = 17.75, p < 0.0001), as well as SOCIAL RELATION and ENVIRONMENT$_{(s\ vs.\ u)}$ ($\hat{\beta}$ = -0.69, $\chi^2(1)$ = 10.44, p = 0.001).

Due to the significant 2-way interactions, we conducted two sub-analyses, splitting the data into public and private conditions. The output of the models are depicted in Table 3.

The model with random item intercepts fit best the data of the public conditions. The results revealed a significant main effect of AGE in that the higher the age, the lower is the rating ($\hat{\beta}$ = -0.02, $\chi^2(1)$ = 20.68, p < 0.0001). The main effect GENDER and ENVIRONMENT$_{(su\ vs.\ u)}$ did not show any significance.

The model with random item intercepts fit best the data of the private conditions. The results revealed a significant main effect of GENDER, in that females rated lower than males ($\hat{\beta}$ = -0.54, $\chi^2(1)$ = 41.05, p < 0.0001). The main effect of ENVIRONMENT$_{(su\ vs.\ u)}$ turned out significant in that people who grew up in a suburban environment provided lower ratings than people who grew up in an urban environment ($\hat{\beta}$ = -0.89, $\chi^2(1)$ = 33.53, p<0.0001). The main effect AGE did not show any significance.

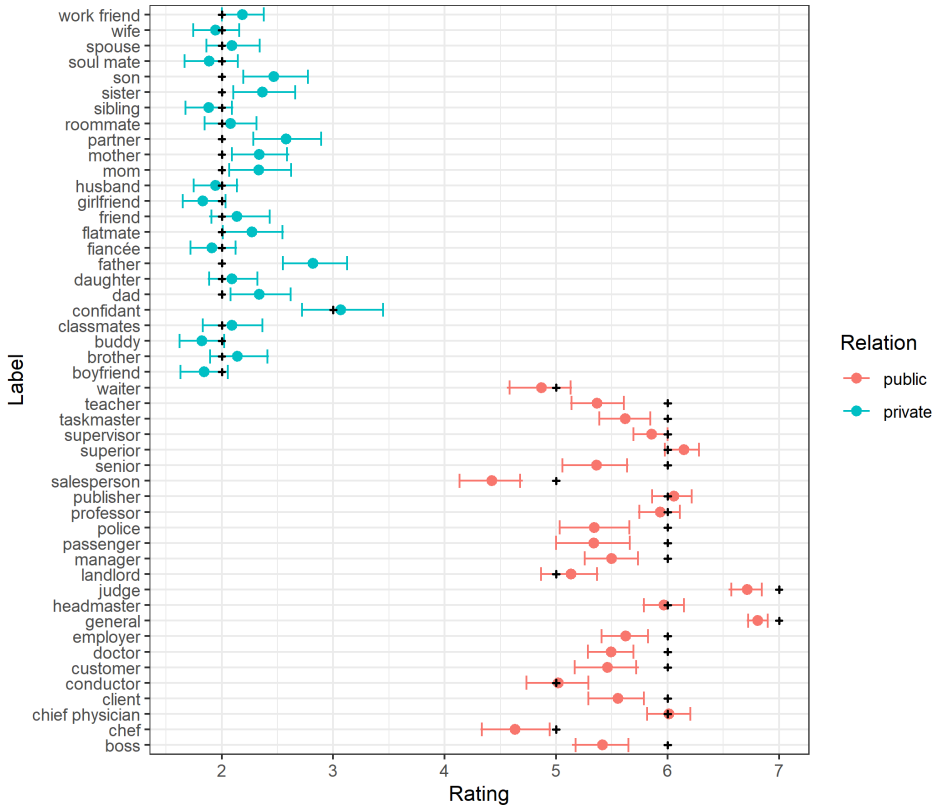Figure 1: Means, standard error, and medians of the formality rating (7-point Likert scale, 1 = Certainly not, 7 = Certainly yes) from the American data set. The x-axis depicts the rating and the y-axis the social relation manipulation. Coloured dots represent the mean ratings, blue shows the public conditions and red the private ones. Bars depict the standard error of the means. Crosses show the median rating.

Table 2: Output of the main analysis with the American data set.

Fixed effects			Model comparison	
Effect	$\hat{\beta}$	SE	$\chi^2(1)$	p-value
SOCIAL RELATION (SR)	-5.12	0.23	521.66	<0.0001
AGE	-0.01	0.01	7.56	0.006
GENDER	-0.29	0.06	24.78	<0.0001
ENVIRONMENT$_{(su\ vs.\ u)}$	-0.53	0.11	24.73	<0.0001
ENVIRONMENT$_{(su\ vs.\ r)}$	0.11	0.10	1.29	0.26
EDUCATION$_{(c\ vs.\ hs)}$	-0.16	0.10	2.83	0.09
EDUCATION$_{(c\ vs.\ g)}$	-0.04	0.11	0.15	0.70
2-way interaction				
SR*AGE	0.02	0.01	9.26	0.002
SR*GENDER	-0.50	0.12	17.75	<0.0001
SR*ENVIRONMENT$_{(su\ vs.\ u)}$	-0.69	0.21	10.44	0.001
SR*ENVIRONMENT$_{(su\ vs.\ r)}$	-0.12	0.19	0.38	0.54
SR*EDUCATION$_{(c\ vs.\ hs)}$	0.28	0.19	2.14	0.14
SR*EDUCATION$_{(c\ vs.\ g)}$	0.21	0.23	0.86	0.35

```
clmm(rating ~ SOCIAL_RELATION + AGE + GENDER + ENV_SU_U
    + ENV_SU_R + EDU_C_HS + EDU_C_G + SOCIAL_RELATION:AGE
    + SOCIAL_RELATION:GENDER + SOCIAL_RELATION:ENV_SU_U
    + SOCIAL_RELATION:ENV_SU_R + SOCIAL_RELATION:EDU_C_HS
    + SOCIAL_RELATION:EDU_C_G + (1 | item), link=logit)
```

Thus, the SOCIAL RELATION and AGE interaction in the main analysis is driven by the significant main effect of AGE in the public condition which lacks in the private condition. The interaction of SOCIAL RELATION and GENDER is due to the significant effects of GENDER in the private condition, which lack in the public relation, and the same is true for the interaction of SOCIAL RELATION and ENVIRONMENT$_{(su\ vs.\ u)}$.

Table 3: Output of the sub-analyses of the American data set.

Fixed effects			Model comparison	
Effect	$\hat{\beta}$	SE	$\chi^2(1)$	p-value
Model 1 – Public conditions				
AGE	-0.02	0.01	20.68	<0.0001
GENDER	-0.05	0.09	0.30	0.58
ENVIRONMENT$_{(su\ vs.\ u)}$	-0.21	0.16	1.85	0.17
ENVIRONMENT$_{(su\ vs.\ r)}$	0.19	0.14	1.89	0.17
EDUCATION$_{(c\ vs.\ hs)}$	-0.33	0.14	5.92	0.02
EDUCATION$_{(c\ vs.\ g)}$	-0.18	0.16	1.17	0.28

```
clmm(rating ~ 1 + AGE + GENDER + ENV_SU_U + ENV_SU_R
    + EDU_C_HS + EDU_C_G + (1 | item), link=logit)
```

	$\hat{\beta}$	SE	$\chi^2(1)$	p-value
Model 2 – Private conditions				
AGE	0.01	0.01	0.11	0.74
GENDER	-0.54	0.09	41.05	<0.0001
ENVIRONMENT$_{(su\ vs.\ u)}$	-0.89	0.15	33.53	<0.0001
ENVIRONMENT$_{(su\ vs.\ r)}$	0.07	0.14	0.23	0.63
EDUCATION$_{(c\ vs.\ hs)}$	0.02	0.14	0.02	0.89
EDUCATION$_{(c\ vs.\ g)}$	0.01	0.16	0.01	0.97

```
clmm(rating ~ 1 + AGE + GENDER + ENV_SU_U + ENV_SU_R
    + EDU_C_HS + EDU_C_G + (1 | item), link=logit)
```

2.3 Experiment 2: British English

This study was conducted simultaneously to the American study and the experimental set-up and data handling were identical to the ones described in Section 2.1. The mere differences concern spelling and label adjustments to British English (e. g. *headmaster* instead of *principal*). Thus, in the Section 2.3.1, we detail the demographic and linguistic background of the participants.

2.3.1 Participants

We recruited 100 British English speakers (female: 49, male: 51, non-binary: 0). We removed one participant who indicated to be non-native to English. All participants matched the inclusion age range. Based on the results of the comprehension question, we removed the data of 16 participants.

The remaining data consisted of 83 participants with the mean age of 35.1 years (SD=11.5) ranging from 19 to 64 years. 20 participants reported to be a dialect speaker (i. e. Birmingham, Geordie, Glaswegian, Scottish, Teesside, Thames Estuary, West Country, and Yorkshire British English dialect). Almost half (46 %) of the participants have a college degree, 39 % obtained a graduate diploma, and 16 % finished high school. The majority (54 %) of the participants grew up in a suburban environment, while 31 % were raised in an urban and 14 % in a rural environment.

2.4 Results

194 entire trails were removed from the data set based on the reading times. The descriptive statistics of the remaining data across the conditions are shown in Table 4. Figure 2 depicts the means, medians, and standard deviations as well as errors of each label. Means and medians of the public condition range from 5.3 to 6.8, while those of the private condition range from 1.8 to 2.5. In the public condition, most means and medians received the rating 5 or higher. The labels *chef* (mean = 4.41, SD = 1.39, median = 5), *conductor* (mean = 4.84, SD = 1.20, median = 5), *salesperson* (mean = 4.00, SD = 1.45, median = 4), *teacher* (mean = 4.89, SD = 1.52, median = 5), and *waiter* (mean = 4.89 SD = 1.52, median = 5) received the lowest means and medians. *General* (mean = 6.67, SD = 0.65, median = 7) and *judge* (mean = 6.78, SD = 0.55, median = 7) received the highest mean and median. In the private condition, most means and medians range from 1.8 to 2.5 with the exception of the label *confidant* (mean = 3.23, SD = 1.66, median = 3).

Table 4: Descriptive statistics of the ratings of the British data set.

Social relation	N	Median	Mean	SD	SE
public	1923	6	5.45	1.27	0.03
private	1867	2	2.09	1.08	0.02

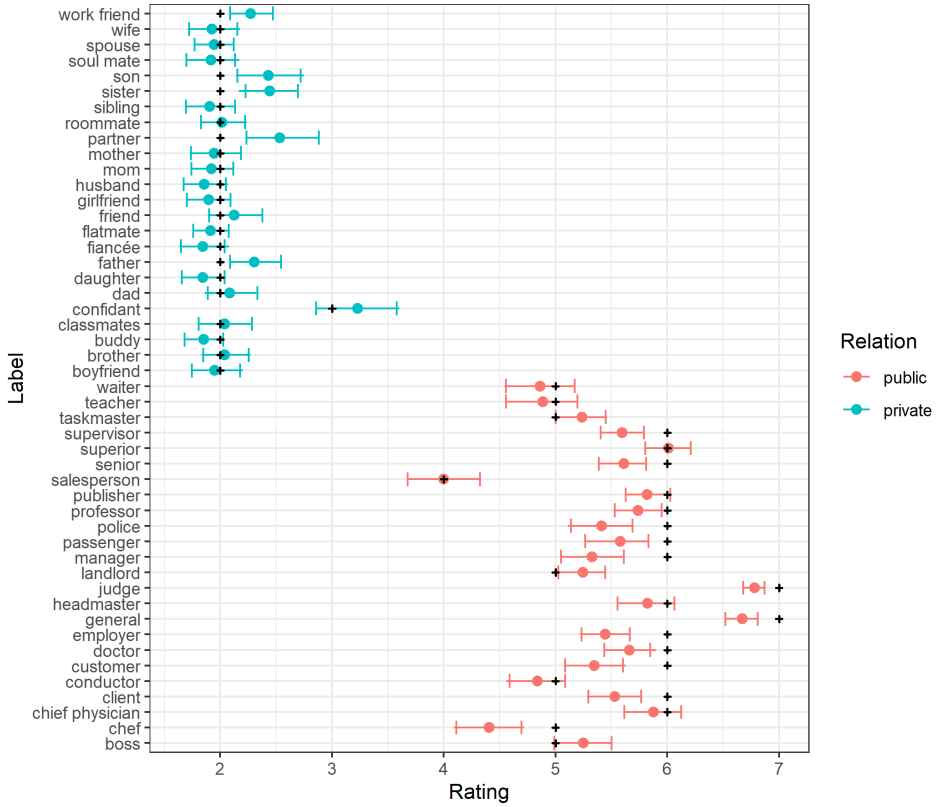Figure 2: Means, standard error, and medians of the formality rating (7-point Likert scale, 1 = Certainly not, 7 = Certainly yes) from the British data set. The x-axis depicts the rating and the y-axis the social relation manipulation. Coloured dots represent the mean ratings, blue shows the public conditions and red the private ones. Bars depict the standard error of the means. Crosses show the median rating.

For the main analysis, the model with random item intercepts[6] fit best the data. The output of the model is depicted in Table 5. The results showed a significant main effect of SOCIAL RELATION in that private conditions received lower formality ratings than public conditions ($\hat{\beta}$ = -3.94, $\chi^2(1)$ = 281.62, p < 0.0001). AGE showed a significant main effect in that the higher the age, the lower the rating ($\hat{\beta}$ = -0.02, $\chi^2(1)$ = 31.47, p < 0.0001). GENDER turned out significant in that females provided lower ratings than males ($\hat{\beta}$ = -0.17, $\chi^2(1)$ = 6.84, p = 0.009). Furthermore, the following 2-way interaction turned out significant: SOCIAL RELATION and AGE ($\hat{\beta}$ = -0.03, $\chi^2(1)$ = 26.82, p < 0.0001), SOCIAL RELATION and ENVIRONMENT$_{(su\ vs.\ u)}$ ($\hat{\beta}$ = -0.95, $\chi^2(1)$ = 22.70, p < 0.0001), and SOCIAL RELATION and ENVIRONMENT$_{(su\ vs.\ r)}$ ($\hat{\beta}$ = 1.40, $\chi^2(1)$ = 30.21, p < 0.0001).

Due to the significant 2-way interactions, we conducted two sub-analyses, splitting the data into public and private conditions. The output of the models are depicted in Table 6. The model with random item intercepts slopes fit best the data of the public conditions. The results revealed a significant main effect of ENVIRONMENT$_{(su\ vs.\ u)}$ in that participants who grew up in a suburban environment showed higher ratings than those who grew up in an urban environment ($\hat{\beta}$ = 0.37, $\chi^2(1)$ = 7.05, p = 0.008). ENVIRONMENT$_{(su\ vs.\ r)}$ showed a significant main effect in that participants who grew up in a suburban environment showed lower ratings than those who grew up in a rural environment ($\hat{\beta}$ = -0.68, $\chi^2(1)$ = 14.48, p = 0.0001). AGE did not show a significant effect.

The model with random item intercepts and AGE and GENDER slopes fit best the data of the private conditions. The results revealed a significant main effect of AGE in that the higher the age, the lower is the rating ($\hat{\beta}$ = -0.03, $\chi^2(1)$ = 18.74, p < 0.0001). ENVIRONMENT$_{(su\ vs.\ u)}$ showed a significant main effect in that participants who grew up in a suburban environment showed lower ratings than those who grew up in an urban environment ($\hat{\beta}$ = -0.62, $\chi^2(1)$ = 57.38, p < 0.0001). ENVIRONMENT$_{(su\ vs.\ r)}$ showed a significant main effect in that participants who grew up in a suburban environment showed higher ratings than those who grew up in an rural environment ($\hat{\beta}$ = 0.66, $\chi^2(1)$ = 45.47, p < 0.0001).

Thus, the SOCIAL RELATION and AGE interaction in the main analysis is driven by the significant main effect of AGE in the private condition, which lacks in the public condition. The interaction of SOCIAL RELATION and ENVIRONMENT$_{(su\ vs.\ u)}$ as well as ENVIRONMENT$_{(su\ vs.\ r)}$ is due to a crossover effect of ENVIRONMENT in public vs. private conditions.

---

[6]Even though a more complex model fit the data better, it caused various computational issues in the generation of the null models. Thus, we reduced the model again and ended up with the simplest one.

Table 5: Output of the main analysis with the British data set.

Fixed effects				Model comparison
Effect	$\hat{\beta}$	SE	$\chi^2(1)$	p-value
SOCIAL RELATION (SR)	-3.94	0.24	281.62	<0.0001
AGE	-0.02	0.01	31.47	<0.0001
GENDER	-0.17	0.07	6.84	0.009
ENVIRONMENT$_{(su\ vs.\ u)}$	-0.13	0.10	1.78	0.18
ENVIRONMENT$_{(su\ vs.\ r)}$	0.16	0.13	1.58	0.21
EDUCATION$_{(c\ vs.\ hs)}$	-0.14	0.12	1.44	0.23
EDUCATION$_{(c\ vs.\ g)}$	0.16	0.09	3.14	0.08
2-way interaction				
SR*AGE	-0.03	0.01	26.82	<0.0001
SR*GENDER	0.13	0.13	1.05	0.31
SR*ENVIRONMENT$_{(su\ vs.\ u)}$	-0.95	0.20	22.70	<0.0001
SR*ENVIRONMENT$_{(su\ vs.\ r)}$	1.40	0.26	30.21	<0.0001
SR*EDUCATION$_{(c\ vs.\ hs)}$	-0.39	0.23	2.78	0.10
SR*EDUCATION$_{(c\ vs.\ g)}$	0.02	0.19	0.01	0.91

```
clmm(rating ~ SOCIAL_RELATION + AGE + GENDER + ENV_SU_U
    + ENV_SU_R + EDU_C_HS + EDU_C_G + SOCIAL_RELATION:AGE
    + SOCIAL_RELATION:GENDER + SOCIAL_RELATION:ENV_SU_U
    + SOCIAL_RELATION:ENV_SU_R + SOCIAL_RELATION:EDU_C_HS
    + SOCIAL_RELATION:EDU_C_G + (1 + | item), link=logit)
```

3 Discussion

We conducted two rating studies with American and British native speakers to measure the degree of formality in settings resulting from different social relations. We thereby manipulated on labels associated with the public and private sector, such as *boss* vs. *mother*, and asked participants to predict the formality

Table 6: Output of the sub-analyses of the British data set.

Fixed effects				Model comparison
Effect	$\hat{\beta}$	SE	$\chi^2(1)$	p-value
Model 1 – Public conditions				
AGE	-0.01	0.01	0.66	0.42
GENDER	-0.26	0.09	8.21	0.004
ENVIRONMENT$_{(su\ vs.\ u)}$	0.37	0.14	7.05	0.008
ENVIRONMENT$_{(su\ vs.\ r)}$	-0.68	0.18	14.48	0.0001
EDUCATION$_{(c\ vs.\ hs)}$	-0.04	0.16	0.05	0.83
EDUCATION$_{(c\ vs.\ g)}$	0.27	0.13	4.07	0.04

```
clmm(rating ~ 1 + AGE + GENDER + ENV_SU_U + ENV_SU_R
   + EDU_C_HS + EDU_C_G + (1 | item), link=logit)
```

Model 2 – Private conditions				
AGE	-0.03	0.01	18.74	<0.0001
GENDER	-0.16	0.15	6.91	0.009
ENVIRONMENT$_{(su\ vs.\ u)}$	-0.62	0.15	57.38	<0.0001
ENVIRONMENT$_{(su\ vs.\ r)}$	0.66	0.19	45.47	<0.0001
EDUCATION$_{(c\ vs.\ hs)}$	-0.30	0.17	33.55	<0.0001
EDUCATION$_{(c\ vs.\ g)}$	0.35	0.14	41.66	<0.0001

```
clmm(rating ~ 1 + AGE + GENDER + ENV_SU_U + ENV_SU_R
   + EDU_C_HS + EDU_C_G + (1 + AGE
   + GENDER| item), link=logit)
```

level of the next utterance. We analysed our data with regards to the social relations while including demographic information of participants; Table 7 shows the summary of the relevant results.

Our results showed that the manipulation of social relations is sufficient to suggest a clear division between formal vs. informal settings. In both the Amer-

Table 7: Summary of the most relevant results of Experiment 1 (American English) and Experiment 2 (British English). The symbols "<" and ">" indicate a statistically significant difference in the respective direction. The symbol "↑" indicates increasing and "↓" decreasing the value to the right of the symbol. The symbol "—" shows a lack of significant effect. "*" indicates that this effect was not important for the interaction resolution.

		Effect	American English	British English
Main analysis		Social relation	private < public	private < public
Sub-analyses	*Public*	Age	↑ age ↓ rating	—
		Gender	—	*
		Env$_{su\ vs.\ u}$	—	su > u
		Env$_{su\ vs.\ r}$	*	su < r
	Private	Age	—	↑ age ↓ rating
		Gender	female < male	*
		Env$_{su\ vs.\ u}$	su < u	su < u
		Env$_{su\ vs.\ r}$	*	su > r

ican and British context, we found strong evidence that labels associated with the public sector set up formal settings, whereas labels associated with intimate relations set up informal settings. This confirmed our hypotheses.

Since formality is an important dimension for the study of register, these findings are valuable for the manipulation of experimental approaches, which aim at investigate phenomena in formal vs. informal contexts. One example for the usage of social relations as manipulation was a rating study investigating the register sensitivity of negation and polarity phenomena, such as negative concord and negative polarity items (Rotter & Liu in preparation). The social relation manipulation served to set up a context in which functional equivalent constructions from different varieties of English (e. g. *I did not see anybody/nobody* in standard vs. non-standard English) were rated with respect to the appropriateness of the utterance. The aim of such case studies is to tackle the interplay between grammatical and register knowledge and to understand language use in general.

Besides the overall effects of the public vs. private relation, we also compared the rating of certain labels. In the American and British data sets, the means and

medians of the private condition are approximately similar except for the label *confidant*; it received the highest mean and median, yet its ratings were still below the neutral point (4). Even though the label denotes a concept of intimacy, the lexical item *confidant* might be of higher formality levels. Possibly the low frequency of this word has an influence on the ratings as well. This word exemplifies the variance of parameters influencing the degree of formality.

In the public condition, three sections of different formality degrees emerge cross-culturally. Labels, e. g. *chef*, *conductor*, and *waiter* were rated with the lowest formality degree of 5, while most labels received approximately a rating of 6. Labels, e. g. *general* and *judge*, received approximately the maximal formality degree of 7. Such levels could emerge through different, intertwined social dimensions, such as hierarchy or the social persona associated to the specific label (Agha 2007, Irvine 1979). For instance, the label *judge* describes the concept of a high ranked public officer associated with the persona of authority, while the label *waiter* refers to the staff in gastronomy with a host persona. Thus, interlocutors might also have different goals of communication when addressing either persona, e. g. asking for orders in a military context vs. asking for cutlery during dinner at a restaurant. We will leave the identification of specific parameters to future work.

Moreover, almost all of the used labels suggest a stereotypical male role (in British English: Misersky et al. 2013). A future study manipulating on the gender of the label by either adjusting the label, e. g. *boss – female boss* or by adding a pronoun in the short story, could show the impact of the perceived gender onto the formality level.[7]

Furthermore, in both data sets, we found evidence that the age, gender, and environment the participants grew up in predicted the overall rating. The 2-way interaction of social relation and the differences between suburban vs. urban environment revealed that cross culturally, people from suburban environments provided lower ratings in the private relation condition than people from urban environments.

There were further differences in the data sets: (1) Concerning the effect of ENVIRONMENT, in the British data set, suburban vs. urban also showed an effect in the public condition with the reverse direction: People who grew up in suburban environments provided higher ratings than people from urban environments. Furthermore, suburban vs. rural showed effects in only the British data set. In the public condition, people growing up in suburban environments provided lower ratings than people growing up in rural environments. In the private condition, the effect was the reverse with people growing up in suburban

---

[7]We thank the anonymous reviewer for pointing out the gender influence.

environments providing higher ratings than people growing up in rural areas. (2) Concerning GENDER, in the American data set, we found that gender had an influence on the rating in private relation conditions in that women provided lower formality levels than men. (3) Concerning AGE, in the American data set, age was found as having an influence on the ratings of the public conditions. The finding suggests that the higher the age, the lower is the rating, i. e. lower formal degrees, in the public condition. In the British data set, age had a similar effect in the private condition; here, participants with higher ages provided even lower formality degrees. (1)-(3) indicate that the background of participants have an influence of the perceived formality degrees in certain conditions. However, identifying the source of these effects seems challenging given that our studies' main focus was the identification of situational parameters, i. e. social relations of interlocutors. We will exemplify possible explanations with the effect of age in the American data set. Learning the concept of formality is, similar to registers, a dynamic process and continues across the lifespan (Agha 1998). Our finding could hint at such a development in that younger American adults attribute higher formality ratings for public labels than older adults; the reasons for this effect can, however, be manifold: First, formal speech can be associated with the open acknowledgement of social rank differences (Irvine 1979). While the social rank differences to a boss, doctor, or professor in the early 20s might be large, it could decrease with age, given that people gain higher job titles with more experience. Second, another possibility is that in many cultures, age is an attribute that requires respect, thus older adults are approached by younger adults with higher degrees of formality. Older adults in turn might be in a position to address the interlocutor with lower formality degrees. Lastly, our participants' mean age in our studies was 35 years with a restricted age range of 18 to 65 years. Even though the age distribution was approximately normal, our sample does not represent the full range of ages. A more focused and controlled study comparing the ratings of young with elderly adults would identify potential differences. However, the tendencies that the older the participant, the lower the rating, could help in future studies to understand ratings including formality manipulations.

The findings (1)-(3) represent a first glimpse at potential demographic influences on the evaluation of formality degrees by language comprehenders, the exact nature of these effects have to be subject of further research as the inclusion of demographic information was purely explorative. Our studies, however, showed reliably that a situational parameter, i. e. social relation, in a controlled setting suggests formal or informal contexts for language comprehenders.

4 Conclusion

In this paper, we investigated the connection between formality and social relations in American and British English participants. We hypothesised that labels of private social relations are associated with lower degrees of formality than labels of public social relations. The results strongly confirmed this hypothesis in both data sets. Our studies are among the first showing that social relations are sufficient to evoke formal or informal settings in two language communities.

Before closing the paper, we would like to briefly discuss potential limits of the study. As discussed in the beginning of the paper, the modality of communication (e. g. written text vs. oral speech) has an impact on the perceived formality level. In our study, we targeted oral speeches, i. e. the upcoming utterance of the protagonist. However, this study was conducted with written sentences, which could have an influence on the overall perceived level of formality. Furthermore, our study tested two closely related language communities of English, thus, whether and to what extent the results apply to other language and cultural communities needs to be further investigated. With these factors taken into consideration, we used a variety of private vs. public social relations in the strictly controlled contexts. The results do not only provide evidence for the predictive role of interlocutor relation on the conversational dynamics, but they also have implications and utilities for experimental manipulations in register research. For example, register-related experiments considering social relations can select a subset of relations from Table 8 and 9 (in Appendix), and use their ratings as predictors.

Funding

Data availability statement

The data sets generated for these studies can be found in the osf repository: https://osf.io/7buzr/.

References

Agha, Asif. 1998. Stereotypes and registers of honorific language. *Language in Society* 27(2). 151–193. DOI: 10.1017/s0047404500019849.

Agha, Asif. 2007. *Language and social relations*. Cambridge: Cambridge University Press.

Atkinson, J. Maxwell. 1982. Understanding formality: The categorization and production of 'formal' interaction. *The British Journal of Sociology* 33(1). 86. DOI: 10.2307/589338.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & Harald Baayen. 2015. Parsimonious Mixed Models. *arXiv* 1506. DOI: 10.48550/ARXIV.1506.04967.

Berruto, Gaetano. 2010. 13. Identifing dimensions of linguistic variation in a language space. In Peter Auer & Jürgen Erich Schmidt (eds.), *Volume 1 Theories and Methods: An International Handbook of Linguistic Variation*, 226–241. Berlin: De Gruyter Mouton. DOI: doi:10.1515/9783110220278.226.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, Douglas. 2009. Multi-dimensional approaches. In Anke Lüdeling & Merja Kytö (eds.), *Corpus Linguistics: An International Handbook*, vol. 2, 822–855. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110213881.2.822.

Biber, Douglas & Edward Finegan. 1994. *Sociolinguistic perspectives on register*. English. New York, NY: Oxford University Press.

Biber, Douglas, Randi Reppen & Susan Conrad. 2007. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Brown, Penelope. 2015. Politeness and language. *International Encyclopedia of the Social & Behavioral Sciences*. 326–330. DOI: 10.1016/b978-0-08-097086-8.53072-4.

Christensen, Rune Haubo Bojesen. 2011. *Analysis of ordinal data with cumulative link models — estimation with the r-package ordinal*. https://www.people.vcu.edu/~dbandyop/BIOS625/CLM_R.pdf.

Christensen, Rune Haubo Bojesen. 2019. *Ordinal—regression models for ordinal data*. R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal.

Creber, Clare & Howard Giles. 1983. Social context and language attitudes: the role of formality-informality of the setting. *Language Sciences* 5(2). 155–161. DOI: 10.1016/s0388-0001(83)80020-5.

Ferguson, Charles A. 1959. Diglossia. *WORD* 15(2). 325–340. DOI: 10.1080/00437956.1959.11659702.

Halliday, Michael Alexander Kirkwood. 1964. The users and uses of language. *Language and Society* 10. DOI: 10.5040/9781474211901.ch-001.

Heylighen, Francis & Jean-Marc Dewaele. 1999. *Formality of language: definition, measurement and behavioral determinants*. Internal Report, Center Léo Apostel. Brussels: Free University of Brussels.

Howcroft, David M. & Verena Rieser. 2021. What happens if you treat ordinal ratings as interval data? Human evaluations in NLP are even more underpowered than you think. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. DOI: 10.18653/v1/2021.emnlp-main.703.

Irvine, Judith T. 1979. Formality and informality in communicative events. *American Anthropologist* 81(4). 773–790. DOI: 10.1525/aa.1979.81.4.02a00020.

Koppen, Kim, Mirjam Ernestus & Margot van Mulken. 2019. The influence of social distance on speech behavior: Formality variation in casual speech. *Corpus Linguistics and Linguistic Theory* 15(1). 139–165. DOI: 10.1515/cllt-2016-0056.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.

Liddell, Torrin M. & John K. Kruschke. 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology* 79. 328–348. DOI: 10.1016/j.jesp.2018.08.009.

Lüdeling, Anke, Artemis Alexiadou, Aria Adli, Karin Donhauser, Malte Dreyer, Markus Egg, Anna Helene Feulner, Natalia Gagarina, Wolfgang Hock, Stefanie Jannedy, Frank Kammerzell, Pia Knoeferle, Thomas Krause, Manfred Krifka, Silvia Kutscher, Beate Lütke, Thomas McFadden, Roland Meyer, Christine Mooshammer, Stefan Müller, Katja Maquate, Muriel Norde, Uli Sauerland, Stephanie Solt, Luka Szucsich, Elisabeth Verhoeven, Richard Waltereit, Anne Wolfsgruber & Lars Erik Zeige. 2022. Register: Language Users' Knowledge of Situational- Functional Variation– Frame Text of the First Phase Proposal for the CRC 1412. *Register Aspects of Language in Situation (REALIS)* 1(1). 1–57. DOI: 10.18452/24901.

Misersky, Julia, Pascal M. Gygax, Paolo Canal, Ute Gabriel, Alan Garnham, Friederike Braun, Tania Chiarini, Kjellrun Englund, Adriana Hanulikova, Anton Öttl & et al. 2013. Norms on the gender perception of role nouns in czech, english, french, german, italian, norwegian, and slovak. *Behavior Research Methods* 46(3). 841–871. DOI: 10.3758/s13428-013-0409-z.

Paolillo, John C. 2000. Formalizing formality: An analysis of register variation in Sinhala. *Journal of Linguistics* 36(2). 215–259. DOI: 10.1017/s0022226700008148.

Pescuma, Valentina N., Diana Serova, Julia Lukassek, Antje Sauermann, Roland Schäfer, Aria Adli, Felix Bildhauer, Markus Egg, Kristina Hülk, Aine Ito, Stefanie Jannedy, Valia Kordoni, Milena Kühnast, Silvia Kutscher, Robert Lange, Nico Lehmann, Mingya Liu, Beate Lütke, Katja Maquate, Christine Mooshammer, Vahid Mortezapour, Stefan Müller, Muriel Norde, Elizabeth Pankratz, Angela G. Patarroyo, Ana-Maria Plesca, Camilo Rodríguez-Ronderos, Stephanie Rotter, Uli Sauerland, Gohar Schnelle, Britta Schulte, Gediminas Schüppenhauer, Bianca Maria Sell, Stephanie Solt, Megumi Terada, Dimitra Tsiapou,

Elisabeth Verhoeven, Melanie Weirich, Heike Wiese, Kathy Zaruba, Lars Erik Zeige, Anke Lüdeling & Pia Knoeferle. 2023. Situating language register across the ages, languages, modalities, and cultural aspects: evidence from complementary methods. *Frontiers in Psychology* 13. DOI: 10.3389/fpsyg.2022.964658.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

Rotter, Stephanie & Mingya Liu. in preparation. A register approach to negative concord vs. negative polarity items in English.

RStudio Team. 2022. *RStudio: Integrated Development Environment for R*. RStudio, PBC. Boston, MA. http://www.rstudio.com/.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag. https://ggplot2.tidyverse.org.

Zehr, Jérémy & Florian Schwarz. 2018. *PennController for Internet Based Experiments (IBEX)*. DOI: 10.17605/OSF.IO/MD832.

Appendix

Table 8: Descriptive statistics per label of the American data set.

	label	N	Min	Q1	Median	Q3	Max	Mean	SD	SE
	boss	91	2	5	6	6	7	5.42	1.22	0.13
	chef	87	1	3.5	5	5.5	7	4.63	1.46	0.16
	chief physician	88	2	6	6	7	7	6.01	0.93	0.10
	client	90	2	5	6	6	7	5.56	1.15	0.12
	conductor	90	2	5	5	6	7	5.02	1.30	0.14
	customer	89	2	5	6	6	7	5.46	1.37	0.15
	doctor	91	2	5	6	6	7	5.49	0.98	0.10
	employer	91	2	5	6	6	7	5.63	1.00	0.10
Public social relation	general	90	5	7	7	7	7	6.81	0.45	0.05
	judge	91	3	7	7	7	7	6.71	0.72	0.08
	landlord	90	2	5	5	6	7	5.13	1.20	0.13
	manager	90	2	5	6	6	7	5.50	1.17	0.12
	passenger	86	1	5	6	6	7	5.34	1.52	0.16
	police	90	1	5	6	6.75	7	5.34	1.56	0.16
	principal	89	2	6	6	6	7	5.97	0.86	0.09
	professor	91	2	6	6	6	7	5.93	0.88	0.09
	publisher	88	2	6	6	7	7	6.06	0.82	0.09
	salesperson	90	1	3.25	5	5	7	4.42	1.35	0.14
	senior	91	1	5	6	6	7	5.36	1.37	0.14
	superior	89	2	6	6	7	7	6.15	0.76	0.08
	supervisor	91	3	5	6	6	7	5.86	0.72	0.08
	taskmaster	90	2	5	6	6	7	5.62	1.08	0.11
	teacher	87	1	5	6	6	7	5.37	1.17	0.13
	waiter	91	2	5	5	6	7	4.87	1.38	0.14

Table 8: Descriptive statistics per label of the American data set (Continued).

	label	N	Min	Q1	Median	Q3	Max	Mean	SD	SE
	boyfriend	89	1	1	2	2	5	1.84	1.01	0.11
	brother	85	1	1	2	2	7	2.14	1.25	0.14
	buddy	90	1	1	2	2	6	1.82	0.97	0.10
	classmates	88	1	1	2	2	7	2.09	1.27	0.14
	confidant	89	1	2	3	5	7	3.07	1.75	0.19
	dad	89	1	1	2	3	7	2.34	1.35	0.14
	daughter	88	1	1	2	3	5	2.09	1.02	0.11
Private social relation	father	89	1	2	2	4	7	2.82	1.45	0.15
	fiancée	89	1	1	2	2	6	1.91	0.98	0.10
	flatmate	88	1	2	2	2	7	2.27	1.26	0.13
	friend	88	1	1	2	2	7	2.14	1.21	0.13
	girlfriend	89	1	1	2	2	5	1.83	0.89	0.09
	husband	87	1	1	2	2	5	1.94	0.97	0.10
	mom	87	1	1	2	3	6	2.33	1.32	0.14
	mother	89	1	2	2	3	7	2.34	1.29	0.14
	partner	92	1	2	2	3	7	2.58	1.45	0.15
	roommate	90	1	1	2	2	6	2.08	1.13	0.12
	sibling	86	1	1	2	2	5	1.88	1.01	0.11
	sister	88	1	2	2	3	7	2.36	1.33	0.14
	son	88	1	2	2	3	6	2.47	1.36	0.14
	soul mate	90	1	1	2	2	6	1.89	1.11	0.12
	spouse	88	1	1	2	2	7	2.09	1.16	0.12
	wife	90	1	1	2	2	6	1.94	0.99	0.10
	work friend	87	1	2	2	3	5	2.18	0.95	0.10

Table 9: Descriptive statistics per label of the British data set.

	Label	N	Min	Q1	Median	Q3	Max	Mean	SD	SE
	boss	80	2	5	5	6	7	5.25	1.20	0.13
	chef	79	2	3	5	5.5	7	4.41	1.39	0.16
	chief physician	81	1	5	6	7	7	5.88	1.14	0.13
	client	81	1	5	6	6	7	5.53	1.12	0.12
	conductor	80	2	4	5	6	7	4.84	1.20	0.13
	customer	81	2	5	6	6	7	5.35	1.22	0.14
	doctor	80	2	5	6	6	7	5.66	0.99	0.11
	employer	81	2	5	6	6	7	5.44	1.01	0.11
Public social relation	general	79	4	6.5	7	7	7	6.67	0.65	0.07
	headmaster	79	2	5	6	7	7	5.82	1.14	0.13
	judge	78	5	7	7	7	7	6.78	0.45	0.05
	landlord	81	2	5	5	6	7	5.25	0.97	0.11
	manager	80	2	5	6	6	7	5.32	1.30	0.15
	passenger	83	1	5	6	7	7	5.58	1.33	0.15
	police	80	2	4.75	6	6	7	5.41	1.32	0.15
	professor	77	3	5	6	6	7	5.74	0.95	0.11
	publisher	78	4	5	6	6	7	5.82	0.89	0.10
	salesperson	80	1	3	4	5	6	4.00	1.45	0.16
	senior	80	2	5	6	6	7	5.61	1.00	0.11
	superior	81	3	5	6	7	7	6.01	0.93	0.10
	supervisor	82	3	5	6	6	7	5.60	0.91	0.10
	taskmaster	80	2	5	5	6	7	5.24	1.03	0.12
	teacher	81	1	4	5	6	7	4.89	1.52	0.17
	waiter	81	1	4	5	6	7	4.86	1.40	0.16

Table 9: Descriptive statistics per label of the British data set. (Continued)

	Label	N	Min	Q1	Median	Q3	Max	Mean	SD	SE
	boyfriend	79	1	1	2	2	6	1.95	1.02	0.11
	brother	78	1	1	2	2	6	2.04	0.99	0.11
	buddy	81	1	1	2	2	5	1.85	0.79	0.09
	classmates	78	1	1	2	2	5	2.04	1.10	0.12
	confidant	78	1	2	3	4.75	7	3.23	1.66	0.19
	dad	72	1	1.75	2	2	6	2.08	1.04	0.12
	daughter	76	1	1	2	2	5	1.84	0.80	0.09
	father	79	1	2	2	3	6	2.30	1.08	0.12
	fiancée	77	1	1	2	2	6	1.84	0.93	0.11
Private social relation	flatmate	79	1	1.5	2	2	4	1.91	0.66	0.07
	friend	80	1	1	2	2	6	2.12	1.14	0.13
	girlfriend	77	1	1	2	2	5	1.90	0.84	0.10
	husband	76	1	1	2	2	5	1.86	0.89	0.10
	mom	78	1	1	2	2	5	1.92	0.82	0.09
	mother	76	1	1	2	2	5	1.95	0.95	0.11
	partner	77	1	2	2	3	6	2.53	1.45	0.17
	roommate	76	1	1.75	2	2	5	2.01	0.92	0.11
	sibling	75	1	1	2	2	5	1.91	0.95	0.11
	sister	79	1	2	2	3	6	2.44	1.15	0.13
	son	79	1	2	2	3	7	2.43	1.38	0.16
	soul mate	83	1	1	2	2	6	1.92	1.05	0.12
	spouse	74	1	1	2	2	4	1.95	0.79	0.09
	wife	79	1	1	2	2	6	1.92	1.03	0.12
	work friend	81	1	2	2	2	5	2.27	0.88	0.10