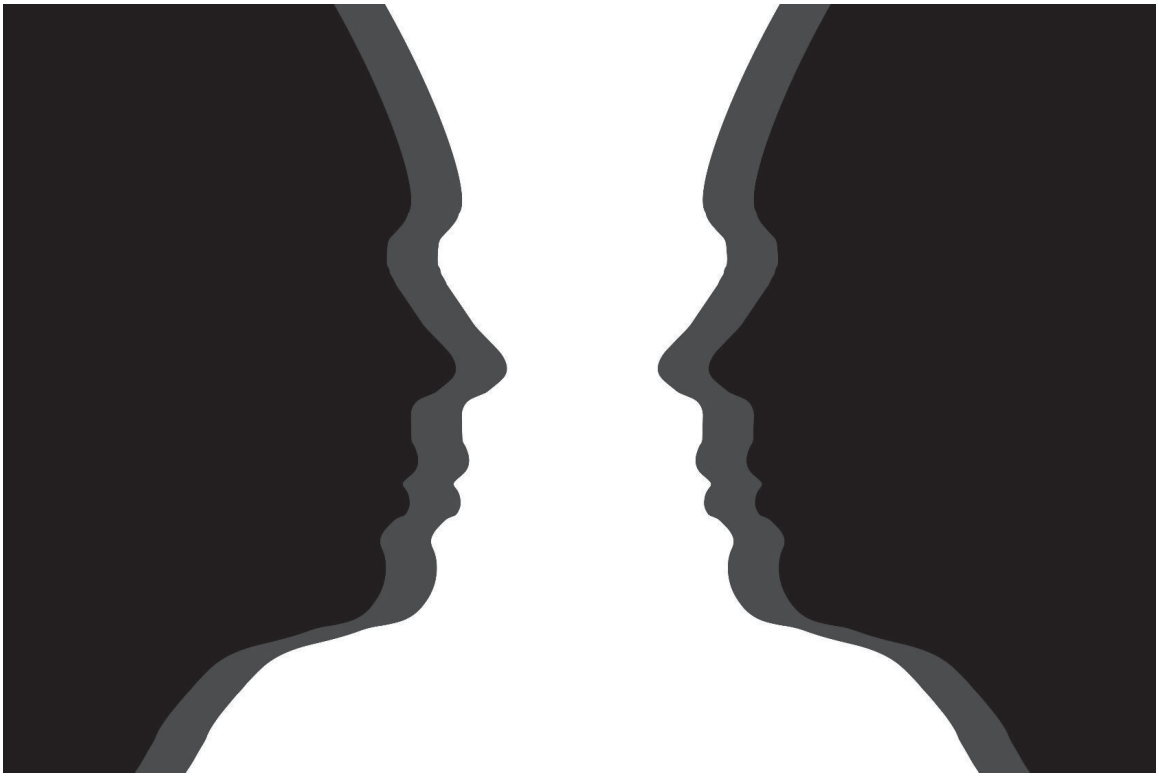
Everyday Diplomacy

Roos, Carla

DOI:
10.33612/diss.230455324

Chapter 1

Overview of the thesis



Overview of the thesis

In our digitizing societies, more and more conversation takes place in online environments, ranging from WhatsApp to Teams, and from e-mail to Twitter. We share our day with our extended family, we congratulate our colleague with her baby, we plan a get together with our friends, we share holiday pictures with our "followers", we file a complaint to a customer service chatbot, and we discuss politics with strangers. But these online interactions, especially the latter, do not always fare well. Indeed, in the context of controversy, online interactions seem to be more prone to escalate into conflict and polarize opinions than discussions held face-to-face (Anderson et al., 2018; Coe et al., 2014; Davis, 2009; Yarchi et al., 2021). Consequently, a lot of research and theorizing has focused on why this is the case. It is virtually impossible, however, to directly compare online with face-to-face (FtF) conversation since there are so many differences between these media that might all (partially) explain their differing proneness to conflict, such as non-verbal cues, anonymity, co-presence, algorithms, etcetera. We reasoned that we can learn more about this between-medium difference by studying what happens *within* online and FtF discussions. How do interaction partners interact and how does this affect their relationship? This is what we aimed to uncover in this thesis by taking a close look at interaction behavior and social perceptions – the social dynamics – within discussions held online and FtF. Our resulting line of research suggests a new way of theorizing about, and doing research into, mediated communication: the social dynamics framework of mediated communication. In Chapter 2, we present this theoretical and methodological framework, and show how it can complement and benefit the existing mediated communication literature. This chapter was written last but forms the opening chapter of this thesis because it provides readers the best understanding of our most up to date thinking and helps contextualize the other chapters in this thesis, both vis à vis each other and vis à vis the broader literature.

The thesis Chapters 3 through 8 are empirical chapters, each consisting of one or multiple studies into online and FtF discussions. We used a broad arsenal of methodologies to give a richer and more complete insight into the (often subtle) processes we studied, which also makes the results more valid and reliable. We performed observational and experimental lab studies with small groups (Chapter 3, 4, 8), analyzed big data from real discussion platforms (Chapter 6), developed new methods for studying online interaction (Chapter 5 and 8), developed and validated a scale to measure conversational experiences (Chapter 7), performed literature reviews spanning diverse fields (Chapter 2 and 7), and developed and tested various interventions (Chapter 8). In each study we complemented the quantitative data with

behavioral observations and/or participant interviews. We used this qualitative data to interpret the quantitative findings. From the behavioral observations, we learned what people actually do and the form their interaction takes: what do people say, how do they say it, and how do others respond? This helped us understand where people's conversational experiences came from. Likewise, from the interviews we learn about the lived experiences of interaction partners: how do they themselves reflect on the quality of their interaction? We also gained insight from actors enacting online and FtF discussions. By in a sense embodying the discussions, they gave us a unique insight into the differing "feel" of online and FtF conversations. We believe that this integrated approach, in which we combined diverse methods and different types of data, enabled us to thoroughly examine the behavioral dynamics and experiences of online and FtF discussions, enriching our understanding of the proneness of online communication to conflict. But beyond that, it helped us to develop a framework to study these, and other between-medium differences, by attending to the within-medium dynamics. We will now shortly describe each empirical chapter in turn.

Chapter 3

In our first empirical study (Chapter 3), we started out exploratively by looking very closely at behavior within text-based online discussions, and comparing this to equivalent FtF discussions. We asked small groups of unacquainted students (N = 108; repeated measures) to discuss via a text-based chat (visually anonymous), a text-based chat with video (no visual anonymity), and FtF. They discussed about issues that they were likely to disagree on, such as closing the national borders to refugees. We performed a content coding on the discussion behaviors and assessed participants' experiences and perceptions in questionnaires.

Behavior-wise, we found that online expression was less ambiguous, or clearer, than FtF expression. For example, whereas FtF people would express their disagreement with sentences like "I ehm maybe sort of do not eh totally agree, I think", online they would rather say "I disagree". Moreover, online interaction is less responsive compared to FtF discussion. This means that participants tended to cross talk more as messages appeared out of order. We know from the linguistic pragmatics literature that these two behaviors - ambiguity and responsiveness - enable people to harmoniously navigate their disagreements in FtF discussions (e.g., Bavelas et al., 1990; Brown & Levinson, 1987; Davis & Perkowitz, 1979). We therefore refer to them as "social regulation techniques", or, in later chapters as "diplomatic skills" or "diplomatic behaviors". The relative lack of diplomatic behavior online seems to be inherent to the limitations of the medium: expression in text is relatively clear and a lack of synchronicity encourages cross-talking. But, this lack of

diplomatic behavior has social consequences. We found that participants experienced more disagreement and less solidarity towards each other online. The lack of diplomatic behavior in a setting where it is expected, might thus lead people to the conclusion there is a problem between them: their opinions might be polarized and they might be in conflict. Interestingly, there were no differences between the visual anonymous (text-only) and non-anonymous (text and video) chats, neither behavior-wise nor experience-wise. This suggests that differences in behavior could not be attributed to a lack of social concern due to visual anonymity (cf. the disinhibition literature, e.g., Suler 2004), but rather to the limited ability to express this social concern via subtle diplomatic skills.

Chapter 4

As the first chapter was mostly explorative and we did not find strong expressions of disagreement or crudeness, in our second empirical study (Chapter 4), we assessed the replicability of these results and tried to study these processes in a more controversial interaction context. We therefore introduced, unbeknown to participants, a research confederate that expressed strong disagreement by advocating a controversial opinion, such as "It is natural that there are few women in the top of business and government". We chose topics on which the participants were likely to agree with each other, and asked the confederate to voice the opposite stance. Besides coding diplomatic behavior and assessing experiences of disagreement and solidarity (like in Chapter 3), we also asked participants for their private opinions on the discussed topics, and to what extent they felt ignored and considered each other disinhibited. We did so to better understand why a lack of responsiveness and a lack of ambiguity might increase experiences of conflict and polarization online.

The study (N = 96; repeated measures) had a similar design to Chapter 1, except for the presence of the confederate, and for the absence of the video-chat condition (since this gave similar results to the chat-only condition). Replicating the results in Chapter 1, we found that online expression was less ambiguous and less responsive than FtF, and that online interaction partners experienced more disagreement and less solidarity. Interestingly, we found that participants did not actually disagree more (or less) online than FtF (as evidenced in their reporting of their private opinions). We further found that participants felt more ignored online and *perceived* each other as more disinhibited, that is, not thoughtful about the social consequences of their behavior. Participants, however, did not consider themselves any more (or less) disinhibited online than FtF, nor did we find evidence for disinhibition in their conversation behaviors. Together, these results suggest that the relative lack of diplomacy online can lead people to perceive polarization and

disinhibition even when there is no "objective" indication for this (in actual opinions, actual behavior). This is important as conflict *perceptions* are likely most consequential for social relationships and behavior (Klofstad et al., 2013; Wojcieszak & Price, 2012).

Taking stock, the first two empirical chapters, by closely examining the microdynamics in conversations (i.e., behavior and its social consequences), offer important new insights into influential assumptions about online discussions: 1) misunderstandings can arise online because people communicate too clearly, rather than too ambiguously as previous literature would suggest (Daft & Lengel, 1986), and 2) people can perceive disinhibition and polarization in conversation due to a misattribution of unintentional non-diplomatic behavior, rather than due to actual disinhibition as previous literature would suggest (Suler, 2004). However, as these conclusions are based on a comparison between text-based online and FtF discussions, we cannot be sure that these social misperceptions are really due to the relative lack of diplomatic behaviors because online and FtF media differ in many more ways.

Chapter 5

In the third empirical study (Chapter 5), we therefore used a new methodology to isolate the influence of responsiveness and ambiguity in shaping perceptions of polarization and conflict. We varied these diplomatic behaviors within media, while keeping other medium differences, such as anonymity, constant. This methodology also enabled us to test whether increasing responsiveness and ambiguity would reduce perceived polarization and conflict. Text-based online and FtF discussions were reproduced in a FtF format (Study 1) and in a text-based chat format (Study 2). In Study 1, we asked a group of actors to enact two online and two FtF discussions from the study reported in Chapter 2. In Study 2, we transcribed these same discussions as if they were all text-based chats. Two new samples of participants ($N_{Study\ 1}$ = 102 and $N_{Study\ 2}$ = 103; both repeated measures) were asked to observe and evaluate these conversations. These observers were not aware of the origin of the conversations, which means that observers in Study 1 thought they saw two original FtF conversations, and observers in Study 2 thought they read two original text-based chats. The results showed that responsiveness was considered indicative of agreement and solidarity in both studies and thus in both medium contexts. But, interestingly, observers considered original FtF conversations more responsive in the FtF context but not in the chat context. Further, independently of the medium context in which it was presented, observers thought interaction partners in the original FtF discussions expressed themselves more ambiguously than those in the original online discussions. The consequences of ambiguity, however, depended on

the medium context: whereas in a FtF context, ambiguity was considered indicative of agreement and solidarity, in a chat context it was perceived as a sign of conflict. This suggests that what counts as diplomacy is medium-dependent.

The interviews with the actors about their (re-)enactment experiences are in line with these conclusions. They thought the online chats were very clear and to-the-point, which disrupted the smooth transition between interaction partners, but also communicated that chatters were confident and at ease. The FtF interactions, on the other hand, were filled with content-less speech, which made conversation very responsive, but also raised the impression that partners felt very tense and uncomfortable. This latter observation might explain why ambiguity does not always work online: observers might have thought that the chatters were so unnaturally ambiguous because they tried to circumvent an imminent escalation of conflict by walking on their toes.

We conclude that the diplomatic behaviors and consequent social dynamics play an important role in misperceptions of polarization, independently of other medium characteristics. In the aggregate, these misperceptions on a microlevel might accumulate into macrolevel societal polarization. To be able to conclude this, however, we need to establish the generalizability of our lab findings to real-life and large-scale online discussions. Indeed, more and more of our everyday online interaction takes place on social media platforms where people communicate to a masspersonal audience rather than small-scale instant chats. Do we also observe variations in diplomatic behavior in social media discussions and does this relate to the effectiveness of social regulation?

Chapter 6

That is what we aimed to uncover in the fourth empirical paper by analyzing discussions on the discussion platform Reddit (Chapter 6). This research was conducted together with prof. dr. Sona Utz from the University of Tübingen in Germany, whom I visited in September and October 2020. We saw in the previous chapter that ambiguity of the FtF type might be interpreted negatively in a chat context. It seems reasonable to suspect, however, that a "milder" form of ambiguity might be profitable in online discussions. Rather than an excessive presence of "ehm" and "sort of", a couple of keywords showing that an opinion is presented as an opinion (e.g., "I think that might not be true") instead of as a fact (e.g., "That is not true") might make disagreeing messages appear less extreme and polarized. This ambiguated disagreement might de-escalate the subsequent discussion by inviting more friendly replies.

To test this, we analyzed Reddit forums ("subreddits") with differing social norms: subreddits known for civil debate (N = 4594), subreddits considered incivil discussion spaces (N = 2126), and subreddits dedicated to social support (N = 1401). We analyzed the comments in these discussions with a validated and widely used automated text analysis program: Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2001; Tausczik & Pennebaker, 2010). With this program, we counted the presence of words signaling ambiguity, such as "think", "believe" "maybe", and words indicating agreement versus disagreement, like "agree", "yes" versus "no", "never". Our outcome variable was the sentiment of the comment represented by the weighted percentage of positive, negative, and neutral emotion words, such as "love", "happy", " f*ck", "asshole".

We found proof for the existence of diplomatic behavior: posters tended to express their disagreements ambiguously. But this did not predict the sentiment in the subsequent discussion. An explorative manual discourse analysis uncovered why this was the case: a lot of discourse lacked dialogue. Posters seemed to be more concerned with sending than with responding and reacting. This makes diplomacy irrelevant. This finding underlines what we found in the lab: online, there is less dialogue because responsiveness is lacking, but since we still expect dialogue, we deduce social conflict. This implies that what appears to be a polarized discussion online might actually be a string of individual statements broadcasted in a shared space. The current study also stresses once again the need to look at microlevel interaction behavior, even in this type of "big data". The usefulness of automated content analysis in studying interaction dynamics is limited.

Chapter 7

In all empirical lab studies (Chapters 3 through 5) and also suggestively in the large-scale observational study (Chapter 6), we saw that the relative lack of responsiveness in online interaction was the strongest and most consistent predictor of interaction partner's perceptions of disagreement and lacking solidarity. The underlying psychological process here appears to be that people feel less heard online: as others are not responding to what they say, people feel ignored, which is detrimental to their social relationship. More generally, feeling heard seems to be of central importance in many domains of life, ranging from romantic relationships and parent-child interactions, to work meetings and filing a complaint with a helpdesk, to therapeutic sessions and citizen-government relations (as evidenced in newspaper articles and popular media, for example, Floer, 2020; Glaser, 2020; Goldberg, 2021, Mohl, 2020). Paradoxically, the scientific literature on feeling heard is scarce and scattered across disciplines. In Chapter 7 we therefore aimed to uncover what it means to feel heard in conversation, and to develop a scale to measure it. We started

with expert interviews and a comprehensive literature review of concepts related to feeling heard, culminating in a conceptualization of feeling heard as consisting of five aspects at two levels of analysis. At the interpersonal level interaction partners feel heard when they have 1) voice, and receive (2) attention, (3) empathy, and (4) respect. At the collective level interaction partners should experience (5) common ground. We accordingly define feeling heard as *the feeling that one's communication is received with attention, empathy, respect, and in a spirit of mutual understanding*. In two population survey studies and a lab study, we subsequently developed and validated a scale to measure this experience: the feeling heard scale (FHS); a concise eight-item scale with good psychometric properties. Specifically, in Study 1 (N = 194), we developed the scale, in Study 2 (N = 1000), we tested its convergent, divergent and predictive validity, and in Study 3 (N = 74; repeated measures), we tested the performance of the scale in a lab experiment.

Feeling heard is a process variable and thus part of the conversational microdynamics we studied before. It is a subjective experience that is influenced by and influences moment-to-moment behaviors (e.g., listening behaviors and conflict intentions, respectively). There are not many scales that assess conversation processes and even less scales that do so at multiple levels of analysis: feeling heard is about me, you, and us, all at the same time. The power of this scale is visible in Study 2, which demonstrated that the FHS was a distinct and powerful predictor of conversation intentions in many different contexts and relationships, ranging from business meetings or fobbing off an advertiser, to intimate conversations with a partner or catching up with friends. In fact, the FHS was the strongest predictor of intentions for conflict behavior among a set of 15 related variables (e.g., acquaintance, dominance, intimacy). Moreover, Study 3 showed that the FHS was highly sensitive to picking up differences in communication setting and explained variance in conversational experiences that other related variables did not. The subtle difference in experiences of *speaking* via text-based communication versus *listening* to text-based communication were uniquely captured by the FHS. In line with Chapter 4, we also found in Study 3 that interaction partners feel less heard when they communicate via text-based chats than via audio-channels. Considering that feeling heard plays an important role in many interpersonal contexts, knowing how to assess and, if needed, increase feeling heard experiences can be crucial.

Chapter 8

As this PhD project was aimed at learning about the reasons behind online susceptibility to polarization, we hoped that our results would also suggest ways to reduce social misunderstandings in online discussions. That is why we designed various interventions in which we tried to manipulate responsiveness and ambiguity

in both online and FtF discussions. Three of these intervention studies are reported in the last chapter of this thesis: Chapter 8. In the first two intervention studies we tried to promote responsiveness and ambiguity in an online chat context, and in the third study we tried to reduce responsiveness and ambiguity in a FtF context. In intervention Study 1 (N = 68; repeated measures), we designed a chat function where participants could see what their interaction partner was typing when they were typing. In intervention Study 2 (N = 74; repeated measures), we designed a new keyboard-like tool that participants could use alongside the regular computer keyboard and allowed them to make the interjecting sounds people frequently use in FtF conversations, e.g., "hmhm", "yes": the VoxBox. In intervention Study 3 (N = 105; repeated measures), we asked participants to start their FtF group discussion by writing down and reading out aloud their opinion on the discussion topic.

In all three studies we were not successful in manipulating responsiveness and ambiguity. We learned more about why this was the case by analyzing participants' behavioral responses to the interventions. These showed that participants tried to compensate for (studies 1 and 2) and/or distance themselves from (studies 2 and 3) the manipulations. This suggests that participants experienced the changes introduced by the interventions as disrupting their conversation and thereby threatening the relationship with their interaction partners. It also suggests that participants (consciously or not) tried to preserve a pleasant conversation and good relationships by falling back on the behavioral norms they associated with the medium. In line with Chapter 5, this suggests that people might have medium-specific expectations for diplomatic behavior. We also find indications that this tendency might backfire in the online context, however. Participants in Study 1 tried to formulate themselves as precisely as possible in order to prevent relational damage, but, paradoxically, their interaction partners interpreted the resulting clear and self-contained comments as evidence of firm conviction.

In sum, this set of intervention studies suggests that the best way to promote harmony and prevent polarization online may not be to change behavior (because people will push back and unintended negative consequences can occur) or to change motivations (because people are already socially motivated), but to improve dialogue. This means that we need to shift focus from the individual interaction partners, and their motivations and behaviors, to them as part of an interacting social unit, in which behaviors and the social consequences of these behaviors are affected by the communication context.

Overarching Comments

Each chapter in this thesis was written as an individual paper. Consequently, the chapters contain some overlap in their descriptions of the theoretical

background, methodology, and conclusions. The second chapter is an overview paper, and the other 6 chapters are empirical papers. At the time of printing this thesis the publication status of the chapters is as follows: Chapters 3, 4, 5 and 8 are published, Chapters 6 and 7 are under review, and Chapter 2 is being prepared for submission.

As a lot of the empirical studies reported in this thesis were exploratory in character, these were not preregistered, but their findings culminated in crucial tests that were preregistered prior to data collection (Study 2 in Chapter 5, Study 2 in Chapter 7, and Study 2 and 3 in Chapter 8). Further, the quantitative data and the coded qualitative data from all empirical chapters is publicly available on Dataverse (except for Chapter 6 where participants could not give informed consent). All the attendant research materials are available on request.