Configuring Audiences: A Case Study of Email Communication

JUSTINE ZHANG*, Cornell University, USA
JAMES W. PENNEBAKER, University of Texas at Austin, USA
SUSAN T. DUMAIS, Microsoft Research, USA
ERIC HORVITZ, Microsoft Research, USA

When people communicate with each other, their choice of what to say is tied to their perceptions of the audience. For many communication channels, people have some ability to explicitly specify their audience members and the different roles they can play. While existing accounts of communication behavior have largely focused on how people tailor the content of their messages, we focus on the *configuring* of the audience as a complementary family of decisions in communication.

We formulate a general description of audience configuration choices, highlighting key aspects of the audience that people could configure to reflect a range of communicative goals. We then illustrate these ideas via a case study of email usage—a realistic domain where audience configuration choices are particularly fine-grained and explicit in how email senders fill the `To` and `Cc` address fields. In a large collection of enterprise emails, we explore how people configure their audiences, finding salient patterns relating a sender's choice of configuration to the types of participants in the email exchange, the content of the message, and the nature of the subsequent interactions. Our formulation and findings show how analyzing audience configurations can enrich and extend existing accounts of communication behavior, and frame research directions on audience configuration decisions in communication and collaboration.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: audience; social interaction; email

ACM Reference Format:
Justine Zhang, James W. Pennebaker, Susan T. Dumais, and Eric Horvitz. 2020. Configuring Audiences: A Case Study of Email Communication. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 62 (May 2020), 26 pages. https://doi.org/10.1145/3392871

1 INTRODUCTION

When people communicate with one another, they have their audience in mind. The nature of the audience is closely coupled with their goals in communicating and their choice of what to say. Prior studies across a variety of settings have demonstrated that accounting for a speaker's beliefs about the audience is key to understanding their communication behaviors [7, 17, 30, 63, 69]. In fact, in many of these settings, speakers have some ability to directly specify and structure their audience.

*The research described in this paper was conducted while the first author was an intern at Microsoft Research.

Authors' addresses: Justine Zhang, Cornell University, USA, jz727@cornell.edu; James W. Pennebaker, University of Texas at Austin, USA, pennebaker@mail.utexas.edu; Susan T. Dumais, Microsoft Research, USA, sdumais@microsoft.com; Eric Horvitz, Microsoft Research, USA, horvitz@microsoft.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.
2573-0142/2020/5-ART62 $15.00
https://doi.org/10.1145/3392871

Consider an illustrative example where this control of the audience is especially fine-grained: calling a meeting. A meeting's initiator can specify exactly who to invite into the meeting room. They can also specify how these participants are organized: whether certain individuals are singled out as facilitators, notetakers, or spokespeople—or even how attendees and chairs are arranged around the room. Further possibilities arise as the meeting progresses; for instance, two participants may quietly branch off into a side discussion. An initiator's choice of *audience configuration*, and a participant's subsequent decision to *reconfigure* the audience, reflects their various goals in the meeting, such as involving relevant stakeholders, calling out responsible parties, or disclosing more targeted information. Indeed, past research has suggested that along with analyzing the discourse within a meeting, examining decisions about the participants can provide insights into how groups perform crucial tasks such as defining an organization's strategy [40].

Audience-centric decisions come to the fore in widely used computer-mediated settings like group chat and email, in which people have the ability to explicitly configure their audiences. For instance, every time someone crafts an email, they specify the recipients they want to include, and how these recipients are to be organized among the `To` and `Cc` address fields. As such, whenever people communicate via email, they must jointly consider their choices of message and audience.

Despite the prevalence and potential importance of audience-centric decisions, existing research on communication behavior has largely focused on how speakers design the content and language of their message. While some of these studies have illustrated that speakers may account for an audience in crafting a message, they have given less consideration to the extent to which speakers can directly configure their audience as well. This gap is particularly salient in settings like email, whose affordances facilitate particularly explicit and fine-grained choices of audience configuration, in addition to choices of message. If speakers must frequently make both types of choices, then our understanding of how they communicate can be enriched by addressing the gap.

In this work, we examine audience configuration decisions and their relation to other aspects of communicating, such as the design of message content. We start by formulating a general description of audience configurations. We draw from prior literature on how speakers account for their audience in various channels, but extend past work by focusing on communication choices that explicitly and directly shape the audience. We highlight several important aspects of the audience that could be configurable: who's included in the audience, how they might be organized into different communicative roles, and how the audience may be modified as the interaction proceeds. We describe how these audience-centric choices feature in a variety of settings, and can reflect various goals and considerations.

To provide an empirical demonstration of our conceptualization, we consider a realistic setting where people are afforded a meaningful level of control in configuring their audience, and where such choices are visible to us as analysts. As a case study, we examine a large corpus of emails sent among employees of a technology firm, noting that in email communication, the criteria of control and visibility are realized in how senders craft the `To` and `Cc` address fields. We find several systematic ways in which types of configurations reflect the senders and recipients involved, the content of the emails, and properties of the subsequent interactions. While each of these aspects has been examined in existing studies of social communication, our findings illustrate how our analyses of audience configurations can clarify and extend these past accounts. We highlight salient distinctions between emails of different configurations, such as in their propensity to receive replies, which would have been missed without a granular view of audience-centric choices. We also find dependencies between audience configurations and characteristics of messages, such as whether recipients' names are mentioned, which suggest that a sender's choice of configuration could convey signals that complement what's relayed in the text.

Drawing on our general formulation and our empirical observations, we propose that audience configurations should be viewed as fully-fledged communicative decisions that are important to account for in studying communication systems. In our case study, we show how we can shed additional light on communication behaviors in a widely-used medium like email. More broadly, we illustrate how accounting for audience configuration decisions could enrich our understanding of day-to-day communication patterns in organizations like the company we examine.

2 BACKGROUND AND RELATED WORK

A broad range of literature has highlighted at least three important and interrelated aspects of communicating a message: the message, its speaker, and its audience. In most communication channels, a speaker can craft the content of their message. In some cases, the speaker is also able to more directly control the audience that the message is directed at. Past work across fields like linguistics and natural language processing has explored ways to characterize message-centric choices. In contrast, our focus is on characterizing audience-centric choices.

In this section, we describe audience configurations as a category of communication decisions, and motivate their study. Drawing on prior theoretical and empirical literature on social communication, we identify key aspects of the audience which are configurable, point to a broad range of settings across which different forms of configuration decisions occur, and suggest various ways in which these decisions can relate to a speaker's communicative goals. We conclude this section by highlighting a gap and opportunity. Our general description suggests that configuration decisions exist in many settings, and can be intricate and meaningful to speakers. However, past studies of social communication have typically focused on the choice of message content, and have not as extensively explored the extent to which speakers can configure their audience.

2.1 Describing audience configurations

We first identify a few key dimensions along which audience-centric decisions can be made. For each of these dimensions, we draw on theoretical and empirical accounts of communicative audiences to suggest their potential importance; we then suggest how a speaker could, in principle, directly specify that aspect of the audience. As such, we clarify what is meant by configuring the audience, and illustrate how configuration decisions can be granular and multifaceted.

First, speakers account for the people who are *included* in the audience: they intend for their messages to be understood by their particular audience members and design these messages accordingly (see [7, 17, 31] for conceptual descriptions and [26, 63, 69, inter alia] for empirical findings). In our motivating example of the meeting, beyond reacting to the composition of a given audience, the initiator A could explicitly list who to include in the meeting: they may invite B and C as relevant stakeholders while excluding others.

Second, speakers may consider how the audience is *organized*. A speaker may intend for a message to be taken up in different ways by different audience members [7, 17, 18]; for instance, A may issue a request to B in front of C. In this sense, we can view the audience as organized into different communicative roles, with B as the primary addressee and C as an auxiliary participant. Alternatively, A may address B and C on equal terms, indifferent to which addressee takes up the request. Again, the meeting's initiator could explicitly specify this organization through mechanisms like assigning roles in the invitation, verbally calling out a participant, or even arranging the chairs in the room [31].

Third, these properties of the audience could be modified in an ongoing interaction [17]. For instance, when responding to A, B might wish to clarify the request or voice a private concern, and hence start a side discussion with A in a corner. Alternatively, they may pull in a new participant who is better equipped to take up the request. As such, B explicitly *reconfigures* the audience.

2.2 Audience configurations in computer-mediated channels

Audience configuration choices exist in varying forms across different communication channels, beyond our hypothetical meeting. These choices are particularly salient in computer-mediated settings: many of these channels offer speakers access to audience members unhindered by constraints like physical co-location, such that the composition of the audience is no longer subject to potential participants being available to meet. Furthermore, across these settings, past studies have suggested numerous ways in which such choices could be meaningful for speakers, in systematically reflecting particular motivations.

Interpersonal communication settings. In mediums like email, instant messaging and group chat, people are usually afforded some ability to specify the audience of their messages. Message-senders in group chat can choose among multiple channels to contact different people [36]; senders in instant messaging and email are generally able to explicitly list their recipients. These settings also contain affordances that allow senders to signal how their audience is organized, for instance by tagging a subset of recipients, or by arranging recipients into different address fields. Additionally, these audiences can be adjusted over the course of the interaction: new participants may be later added to an exchange, while a reply may target only a subset of the original participants.

Many studies of these settings have found that who people include in their audiences reflects social factors like homophily [45, 68] and power relations [54], events such as internal crises or external shocks [5, 24, 70], and underlying motivations such as fostering social support or seeking different perspectives [11, 33]. Other studies have performed focused studies of particular choices of configuration, taking discourse analysis, interview or survey-based approaches. Several such studies have examined cases when senders include a subset of recipients in `Cc` [10, 35, 52, 76], finding that senders use the mechanism to ensure accountability or transparency, facilitate participation from a wider range of stakeholders, or project one's activities and accomplishments to onlookers. The `Cc` designation has also been suggested as a way to manage recipients' attention [46], as people may perceive emails as less important to them if they are placed in `Cc` [39, 72]. Replying to one person versus replying-all [32] could reflect a sender's desire to manage the cohesion of a group, or their judgment of which participants are still relevant in an exchange.

Social media platforms. Many social media sites offer their users various affordances to manage their audience. Privacy settings allow people to restrict who is included in their audience; by tagging someone on Twitter or commenting to another person's Facebook wall in lieu of making an undirected post, people can organize their audience by specifying a primary target for their communication. Further, subsets of the audience could be bracketed out as people move an interaction from a public forum to a private message. Configuration choices in these settings are comparatively coarse-grained: the audience management options provided by platforms are often limited in scope, cumbersome or not well-understood by users [48, 50].

Despite these limitations, audience-centric considerations are nonetheless relevant to users of these platforms. Past work has pointed out various considerations in choosing between presently available options, such as familiarity with the audience or the willingness to self-disclose [27, 29, 51]. Targeting someone's Facebook wall versus issuing a broad posting could reflect a contrast in objectives related to relationship management versus social validation [6]; the option to more finely specify audience members of WeChat postings gives rise to more elaborate concerns like the social cost of making someone feel snubbed by deliberately excluding them [48]. Furthermore, the paucity of privacy setting options has been cited as a source of frustration for users [48, 50], suggesting that people have some intuition for how various audience configurations could be useful, even if the ability to choose these configurations is presently unavailable to them.

2.3 Existing studies of audience configurations

We have suggested that audience configuration decisions broadly recur across many settings and that they can meaningfully reflect various aspects of communicating. As such, we highlight an opportunity: examining how people choose among a potentially rich range of configurations can add to our understanding of social communication. For instance, past work has examined strategy-setting meetings, showing that such considerations around calling a meeting, in addition to what's discussed within the meeting, are informative of how strategies are developed or reinforced in organizations [40]. In this case, jointly considering the choice and arrangement of participants alongside the choice of discourse was fruitful in analyzing a consequential phenomenon.

In general, however, we find that past studies of social communication have asymmetrically focused on how speakers design their messages. Literature on *audience design* has largely dealt with how speakers tailor the content and style of their messages to suit a given audience [7, 41, 58, 63, 85]. Less attention has been paid to understanding how speakers could configure their audience, especially at the granularity exemplified in our hypothetical offline meeting.

This lack of attention partially reflects that configurations may not be fully controllable by the speaker in many of the settings that have been studied. In fact, especially in large-scale social media platforms, speakers necessarily make inferences about the composition of their audience [8, 49, 53, 81]. That configurations are coarse-grained and not fully legible limits the extent to which we can consider them as meaningful choices for speakers; in the absence of explicit records, these choices are also difficult for researchers to precisely analyze.

In settings where people can make more granular and explicit configuration choices, this degree of control over the audience is largely not emphasized; analyses of channels like instant messaging and email have often abstracted out key aspects of the configuration decision. Many of these studies focus on analyzing the content of messages, largely agnostic to the nature of a message's audience [14, 71, 82]. Some of these studies account for the identities of message recipients in modeling large-scale interaction networks [5, 24, 54, 70, 78], or in studying communication with recipients across organizational strata [28, 61] or within teams [56]. However, these studies do not differentiate between the types of audiences that these recipients are part of, hence ignoring factors like the roles to which recipients are designated. Other work has focused on analyzing messages directed at a narrow subset of the possible configurations afforded by a channel, such as Cc-ing [35, 52, 76] and replying versus replying-all [32]; in a similar vein, many studies restrict the bulk of their analyses to dyadic exchanges [25, 44, 55, inter alia]. These methodological approaches allow analysts to cleanly examine a wealth of phenomena in social communication, providing informative accounts of how speakers communicate and how they account for their audience that we have drawn on. However, taken together, these studies do not provide similarly rich accounts of audience configuration choices.

3 CASE STUDY

We now describe an approach to studying audience configuration choices, that accounts for multiple aspects of the decision and a range of underlying social considerations, as outlined in the preceding section. We apply our approach in a case study of how configurations are used in a large corpus of emails. In this setting, we explore potential insights about social communication that a more granular study of audience configurations could yield.

In what follows, we instantiate our general description of audience configurations in the context of email. We highlight qualities of email communications that make it an especially informative case example, and describe the particular dataset we studied. We additionally highlight particular aspects of emailing and of social communication that we will subsequently examine.

3.1 Setting: Email Communication

Several properties of email make the medium a particularly illustrative setting in which to explore configuration choices. As already noted, an email sender can precisely specify their choice of recipients to include on a message, while the Cc mechanism allows them to explicitly organize these recipients. Furthermore, these granular choices are highly salient to email senders: they decide on a configuration every time they start an email thread, while the option to add or remove recipients is available in any subsequent reply.

Email is also an important case example because it is a realistic and widely used venue in which these configuration decisions are made. It is one of the most popular online activities [1, 67], serving as a key medium for information management and communication across the globe [83]. As such, emailing occupies substantial time and attention for senders and recipients [34, 84].

Dataset. We analyze a corpus of emails sent among employees in a now-defunct technology company referred to with the codeword Avocado. The corpus was compiled from the employees' Microsoft Outlook accounts, and is accessible through the Linguistic Data Consortium [59]; it has been used in several prior studies of emailing [2, 71, 82, 87].

The Avocado collection contains a thorough record of day-to-day emailing activity across an organization. The organization is a realistic and particularly rich venue in which to study social behaviors: communications are structured by the specialized positions people have in the organization, the various tasks that employees have to complete in collaboration with others, and the social relationships induced by roles and subdivisions in the company. The dataset's coverage of personnel across the company distinguishes Avocado from other email collections like Enron, which is focused on the upper management [43].

Data processing. In order to ensure that the dataset faithfully reflects emailing activity in Avocado, we took some additional steps to process the collection that was originally released. In particular, we augmented the data with more information about the reply structure within email threads, matched employees to the multiple email addresses they were associated with, and removed recipients who were duplicated in address fields.[1]

To simplify our initial exploration of audience configurations and to address some additional constraints in the data, we restricted our study to a subset of the collection. First, we considered emails sent only among Avocado employees, ignoring threads which involved people outside the organization. Second, we ignored emails sent to mailing lists or that contained recipients in the Bcc field, since the particular recipients of these messages were often not listed in the data. Finally, we removed self-emails and automated messages. These omitted cases present interesting variations on our basic conceptualization of audience configurations, and are worth exploring in future work.

The collection we subsequently analyze contains 168,452 emails sent among 320 employees between 1999 and 2003. Further details about the dataset, the data processing procedure, and our subsetting decisions can be found in the supplementary material.

3.2 Approach

We use our case study to illustrate the broad variety of potential analyses that follow from our general description of audience configurations. In particular, we aim to explore multiple configurable aspects of the audience, described in Section 2.1, as well as a range of aspects of social communication explored in previous work and surveyed in Section 2.2. To this end, we perform a computational analysis of Avocado, which enables us to make comparisons across configuration choices and examine aggregate patterns across the entire corpus.

[1] Code that implements this procedure can be found at https://github.com/tisjune/avocado-data-processing.

First, we seek to address the multifaceted nature of audience configurations by accounting for inclusion, organization, and reconfiguration choices. We analyze two categories of configuration decisions in turn: when starting an email thread, the sender must decide on a *root configuration*—which recipients to include, and how to organize them among the To and Cc fields; when replying to a previous email, a sender may *reconfigure* the audience through adding or removing recipients.

Second, we seek to explore how analyses of audience configurations could add to existing understandings of various aspects of social communication. In particular, prior studies, including of email, have examined the relationships among participants of an exchange, the intentions of speakers, linguistic properties of messages, and ensuing interaction dynamics. In our case study, we build on these existing analyses as departure points, and explore how different choices of configuration are related to each of these aspects. To interpret the findings of our computational approach, we also draw from prior work, which has used close readings of emails or interviews with senders and recipients to provide more nuanced characterizations of emailing behavior.

4 ANALYSIS: ROOT CONFIGURATIONS

We first examine how email senders configure their audience when they *initiate* an email thread (as opposed to replying to an existing email). When creating these initial emails, henceforth *roots*, the sender selects the recipients of their message, and organizes them into the To and Cc fields. Our focus is on analyzing how senders explicitly make these choices, so we distinguish roots from subsequent replies which, by default, are pre-populated with the configuration of the previous email. As such, we restrict the subsequent analyses to the 79,973 roots in the data.

Representing configurations. To facilitate an automated analysis, we consider abstracted representations of configurations. Importantly, we ensure that these representations, while simplistic, reflect both the inclusion and the organization of recipients.

We address the sender's choice of recipients to include by considering the *audience size*, N. Here, we abstract away from specifying *which* recipients are included in the audience, a modeling choice also taken in past studies of audience design [26, 69]. Future work could more finely account for the identities of the people involved, which we briefly explore as well.

We address the sender's organization of recipients by considering the number of recipients in the To field, which we refer to as the To *size*, k (in our data, $k > 0$). As such, there are $N - k$ recipients in the Cc field. We refer to configurations where $k < N$ (at least one recipient is in Cc) as *tiered*, and configurations where $k = N$ (all recipients are in To) as *flat*. In interpreting our subsequent findings, we assume that people in To versus Cc are generally designated to be primary versus auxiliary addressees [52, 60, 76]. However, we note that senders may vary in their use of these address fields; we later take steps to ensure that our findings are not driven by per-sender idiosyncracies. We also note that senders may signal such designations by other means, such as in the text of the message [17, 31]; future work could account for these other sources of signal.

4.1 Distribution of root configurations

We start by describing the distribution of configurations in Avocado, in terms of audience size N and To size k. Here, we show how the breadth of configurations used by senders extends beyond the particular configurations analyzed in prior work.

Figure 1A shows the proportion of root emails per N. We see that over 40% of threads involve multiple recipients, with $N \geq 2$.

Figure 1B breaks down this distribution by how the audience is organized: the color of the N, k-th heatmap entry denotes the proportion of roots with a To size of k, out of all roots with an audience size of N (i.e., the heatmap is column-normalized). Among emails with multiple recipients, we find that two types of configurations are commonly used, discernable as the dark bands in the figure.

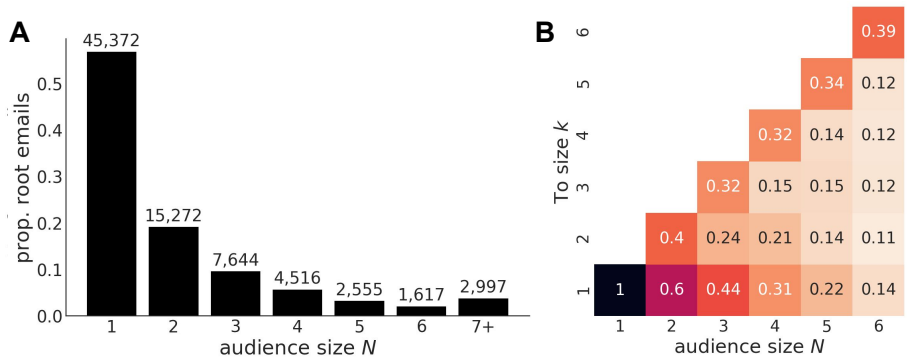

Fig. 1. For *root emails* that initiate a thread: (A) distribution of audience sizes N; the number of emails per N is indicated above each bar. (B) To sizes k per N, over emails with $N \leq 6$ recipients (96% of roots); proportions of roots with To size k out of all roots with audience size N are shown in each cell.

Many prior studies have examined a sender's decision to involve multiple recipients by placing them in the Cc field, thus focusing on tiered configurations [35, 52, 76]. However, we see that in Avocado, senders often place the entire audience in To instead (diagonal), thus opting for a flat organization. For instance, over roots with two recipients ($N = 2$), 40% of emails have a flat configuration with both recipients in To ($k = 2$, corresponding to the top square in the second column from the left). We also note a clear distinction within tiered configurations: typically, only one person is positioned as the primary addressee in To (bottom row), while multiple recipients are "singled out" in To (middle of the heatmap) much less frequently.[2]

When N is large, we draw a potential analogy to posts on social media platforms with large audiences: flat configurations correspond to undirected updates (i.e., to a broad swath of the organization) while tiered configurations correspond to directed communications such as Facebook wall posts [6]. We find that as N increases, flat configurations become more common relative to tiered ones (visually, the diagonal becomes darker than the bottom row after $N = 5$); perhaps fewer situations call for senders to address a particular target in front of a large group of other colleagues.

4.2 Sender and recipient relationships

Next, we explore how configurations reflect the senders and recipients involved. Prior work has shown that the choice to communicate across different types of relationships reflects different considerations: for instance, seeking social support among close relations [20, 29], soliciting novel perspectives from people in different social circles [11, 33], or foregoing interactions with outsiders during a crisis [70]. In a setting like email, a sender can address more than two recipients who potentially span multiple types of relationships. As such, beyond the selection of recipients, they must make decisions about the size and structure of the group of participants, which we would expect to echo the nature of the relationships involved.

To examine these decisions, we use one indicator of sender-recipient relationships accessible in our data: affiliation to divisions in the company. We focus on the two largest divisions, sales and engineering, with 79 and 108 employees respectively.[3] We assume that two people within the same division are more likely to have a closer working relationship than people in different divisions.

[2] As an interpretation of such a case, consider a situation where $N = 3$ and $k = 2$, in which two people are jointly responsible for something, but a third person is also included as a relevant overseer.

[3] Employee affiliations are inferred from job titles in contacts lists included in the corpus.

emails from sales	% tiered	emails from engr.	% tiered
sales → sales (3,906 emails)	59	engr. → engr. (7,723 emails)	61
sales → engr. (337 emails)	43	engr. → sales (166 emails)	41
sales → both (1,145 emails)	68	engr. → both (717 emails)	66

Table 1. Percentages of emails which are tiered (i.e., the recipients are separated into the To and Cc fields, as opposed to all in To), over subsets of multi-recipient root emails, where senders and recipients are in the same or opposite company divisions. Comparing between each pair of rows, Fisher's exact $p < 0.05$.

While coarse, the relative frequency of communications within versus across divisions suggests that this assumption is a valid starting point. 12,690 and 18,062 emails involve people who are all in sales or all in engineering, respectively; in contrast, 4,208 span across the two divisions: 2,466 or 16% of emails from senders in sales include at least one recipient who is in engineering, while 1,742 or 9% of emails from senders in engineering include a recipient from sales.

We decompose the ensuing analyses into two questions: 1) in emailing people within versus across division, when might a sender *include* additional recipients, versus opt for a dyadic message to only one recipient? 2) in emailing multiple recipients in different divisions, when and how does the sender *organize* them into different communicative roles?

Including additional recipients. Prior literature has characterized dyadic communications as uniquely straightforward and intimate, in contrast to exchanges with larger groups [75]. As such, we may expect that emails sent within a division, where senders and recipients are likelier to be familiar with each other, are more often dyadic than emails sent across divisions, in which multiple recipients may be included. For each division, we compare the set of *within-emails* where all senders and recipients are in that division, with *across-emails* where at least one recipient is from the opposite division. Within each set, we compute the percentage of emails which are dyadic. We find that within-emails are indeed more likely to be dyadic than across-emails, for both divisions: 69% of within- versus 40% of across-emails from senders in sales are dyadic; the same proportions for engineering are 57% versus 49% (Fisher's exact $p < 0.001$ for both comparisons).

Organizing multiple recipients. We now focus on emails directed at multiple recipients, examining how these recipients are organized by the sender. In the following analysis, we consider three cases: 1) a sender emails only recipients from their own division; 2) a sender emails only recipients from the opposite division; 3) a sender emails recipients from both divisions. We analyze the configurations used by senders in communicating with each type of audience, in terms of how often senders organize the recipients in a flat configuration with all recipients in To, versus splitting them between To and Cc in a tiered configuration.

Table 1 shows the percentage of emails which are tiered in each case. We see that the share of tiered emails is highest when recipients are from different divisions (i.e., at least one in sales and at least one in engineering, third row), illustrating that the sender's designation of roles tends to echo the presence of different social relationships in the communication. To more precisely examine *which* roles recipients are organized in, we consider the 1,256 emails with recipients from both divisions, and which have tiered configurations. Here, the sender can either designate the recipient in their own division to To and the other to Cc, or the other way around.[4] We see that the *cross-division* recipient is in To more frequently: in 64% and 63% of such emails from a sender in

[4]We ignore the 2.4% of these emails where the designation occurs in both directions.

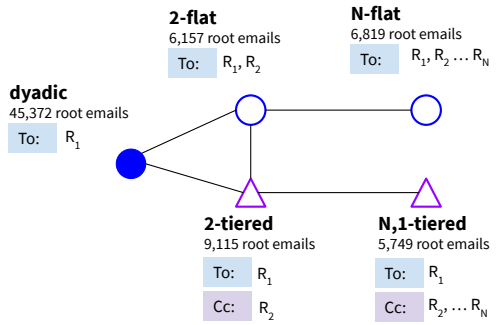

Fig. 2. Visualization of the various pairs of root configurations compared in the analyses in Sections 4.3—4.6. Lines between icons correspond to particular comparisons highlighted in the text and examined using the sender-controlled protocol detailed in Section 4.6.

sales and engineering, respectively. This majority case reflects a previously-observed strategy of including a colleague for additional support or oversight when communicating with a different group, perhaps to facilitate a potentially contentious interaction [10]. In fact, senders tend to designate these colleagues as auxiliary recipients in Cc, thus explicitly distinguishing their intended roles from that of the cross-company addressees in To. This suggests that the configuration of an email's audience may convey or reinforce social cues beyond those contained in its text [61, 77].

We find an interesting contrast in the two cases where all recipients are in the same division: fewer emails have tiered configurations when the recipients are all in the opposite division from the sender, versus when all participants are in the same division (first versus second rows). This perhaps reflects that senders are less prone to making designations among recipients who may be less familiar to them or outside of their chain of command.[5]

4.3 Sender intents

Past work across a variety of settings has examined how people access different audiences to facilitate different communicative intentions [6, 29, 33, inter alia]. Here, we build on these existing analyses to more finely account for how senders can actively configure their audiences to reflect their intentions. In particular, we consider two key types of intents, *requests* and *commitments*, which have been highlighted in the pragmatics literature [4, 74], as well as in specialized taxonomies of emailing intents [19, 21]. To infer these intents from an email, we use classifiers that were previously developed in our organization and applied to the Avocado dataset [82]; the classifiers identify emails with each intent based on the text of the emails.[6] We then compute the proportion of emails of each configuration which contain each type of intent.

In the following analyses, we make comparisons among the most frequent configurations (corresponding to the dark regions in Figure 1B). These comparisons are depicted in Figure 2. We start by examining three minimal cases: *dyadic* (for emails sent to one recipient, $N = k = 1$), *2-flat* (emails sent to two recipients with both in To, $N = k = 2$), and *2-tiered* (emails sent to two recipients

[5]To check that these findings relate to the tiered configuration of the audience, and aren't wholly accounted for by variation in audience size, we repeat the analyses of multi-recipient emails over the 6,994 emails with exactly two recipients (comprising 20% of the multi-recipient emails initially examined), and find that the observed differences still hold in this subset.

[6]To facilitate analyses of email text, we preprocessed emails by removing artifacts like senders' signatures via a set of heuristics. Further details about the text preprocessing procedure and the intent-tagging models that we used are provided in the supplementary material.

	attribute	dyadic vs. 2-flat	2-flat vs. 2-tiered	2-flat vs. N-flat	2-tiered vs. N,1-tiered
(i)	has request (bin.)	↑↑↑	↑↑↑ (*)		↑↑↑
	has commitment (bin.)				
(ii)	has recipient name (bin.)	↓↓↓ (*)	↑↑↑ (*)	↓↓↓	↑↑↑ (*)
(iii)	wordcount	↑↑↑ (*)	↑↑↑ (*)	↑↑↑ (*)	↑↑↑ (*)
	% 1st-person sing.	↓↓↓ (*)			↓↓↓ (*)
	% 2nd-person	↓↓↓ (*)	↑↑↑ (*)		↓↓↓ (*)
	% 1st-person pl.	↑↑↑ (*)	↑↑↑ (*)	↑	↑↑↑ (*)
(iv)	has reply (bin.)	↑↑↑ (*)	↑ (*)	↓↓↓	↑↑↑

Table 2. Significance test statistics on attributes derived from email content (i-iii) along with reply rate (iv). In each column titled A vs. B, attributes of root emails with configuration A are compared to those of root emails with configuration B. Arrow directions indicate whether configuration B is greater in that attribute than A (↑) or vice versa (↓). The number of arrows indicates statistical significance under either the Mann-Whitney U test for real-valued attributes, or Fisher's exact test for binary attributes indicated as (bin.), Bonferroni-corrected in the number of attributes (↑: $p < 0.05$, ↑↑: $p < 0.01$, ↑↑↑: $p < 0.001$). (*) indicates that the comparison is also significant under the sender-controlled protocol described in Section 4.6, indicating consistent effects across different senders (Bonferroni-corrected Wilcoxon signed-rank $p < 0.05$).

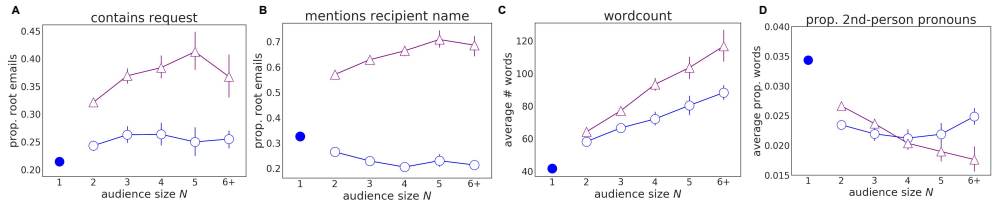

Fig. 3. Averages of various attributes derived from email content, over root emails with different configurations: • denote dyadic emails; ○ denotes N-flat emails for different audience sizes N; △ denotes N, 1-tiered emails. Error bars are bootstrapped 95% confidence intervals.

with one in To and the other in Cc, $N = 2$ and $k = 1$). Comparing emails with dyadic and 2-flat configurations highlights a sender's choice to *include* an additional recipient, while comparing 2-flat and 2-tiered emails highlights the sender's choice to *organize* different recipients across To and Cc. Beyond these base cases, we then compare 2-flat to *N-flat* emails with more recipients in To ($N = k > 2$), and 2-tiered to *N,1-tiered* emails with one recipient in To and additional recipients in Cc ($N > 2$ and $k = 1$). We note that less frequent configurations, which we omit, may reflect interesting communication dynamics which could be taken up in future work.

Statistical significance tests for these comparisons are shown in Table 2(i). Focusing on the relation between configurations and requests, visualized in Figure 3A as the proportion of requests per configuration type, we find that tiered configurations are more often used to convey requests than flat or dyadic ones. We offer two possible interpretations, taken from prior studies of emailing [10, 52, 76]. A sender may explicitly target a recipient in To to take up the request, while Cc-ing overseers to see that the request is carried out. Alternatively, senders may be more prone to making requests as part of collaborations between multiple people; the Cc's could be additional stakeholders who the sender designates as less central to the task, but who might still want to weigh in [46].

4.4 Audience design and linguistic attributes

Past work (surveyed in Section 2) has examined the relation between the content of a message and its audience. Extending these studies, we relate the configuration of an email's audience to its text in a more granular way, accounting for both inclusion and organization choices. We focus on an attribute which directly links a message and its audience: whether the sender mentions a recipient's name [17, 31]. For each configuration, we compute the proportion of emails whose text contains the name of at least one recipient.[7]

Comparisons between configurations are shown in Table 2(ii) and visualized in Figure 3B. We note a striking difference: tiered emails are much more likely to mention recipient names than flat emails. More closely examining *which* recipients are mentioned in the tiered case, we see that recipients in `To` are mentioned more than twice as often as recipients in `Cc`, in 50% versus 22% of tiered emails. The propensity of name-mentioning decreases from dyadic to N-flat emails, perhaps corresponding to addressing a group in its entirety rather than calling out individual members.

These findings suggest several ways in which a message and its audience may be related. First, echoing the audience design literature, senders may tailor their message to suit a given audience, for instance by mentioning the names of audience members. Here, our results underline that audiences may consist of different types of participants, in that name-mentioning clearly reflects some distinction between recipients in `To` and `Cc`.

The fact that email senders can configure, rather than simply react to their audience, gives rise to additional interpretations that go beyond audience design. In particular, audience configurations can bolster or supplement the signal conveyed in message content. In tiered emails, we see that senders tend to call out primary addressees who are already highlighted in the `To` field, while in flat emails, which are perhaps unfocused or uniformly relevant across the N recipients in `To`, these recipients are seldom called out in the text. Further, in over 3/4 of tiered emails, addressees in `Cc` are never mentioned (at least by name), suggesting instead that their intended role in the exchange is to be parsed from the configuration itself.

Additional linguistic attributes. We also explore some other characteristics of the content: the email's length (in number of words), and the use of various personal pronoun categories (e.g., first-person singular), measured as the fraction of words in an email which are of a particular category.[8] These attributes have been used as indicators of group dynamics [15, 42], psychological processes [65], and linguistic style [22, 64]; they comprise a small sample of a wealth of other attributes that future work could explore. Table 2(iii) lists comparisons of these attributes across different configurations; we discuss a selection below.

Figure 3C shows the average lengths of emails with various configurations. We see that length increases with the number of recipients; further, tiered emails are longer than flat ones. This perhaps reflects the relative complexity of configurations [26]—for instance, addressing several people who play different roles in the exchange may require a sender to convey more information.

Figure 3D shows the concentration of second-person pronouns (e.g., *you*) per configuration. We see that these pronouns are used most frequently among dyadic emails, perhaps relating to the focused nature of these communications. Tiered emails contain more second-person pronouns than flat ones, perhaps further reinforcing the sender's focus on their primary addressees; however, the use of these pronouns decreases in N, 1-tiered emails with larger N.

[7] We use some heuristics to infer names and commonly-used nicknames, detailed in the supplementary material.

[8] We use the categories provided in the LIWC lexicon [79], consisting of pronouns and common contractions, e.g., "youre". While this approach would fail to characterize emails not written in English, 96% of emails are in English, as inferred using the python langdetect package [23]; as such, we assume this caveat is relatively minor.

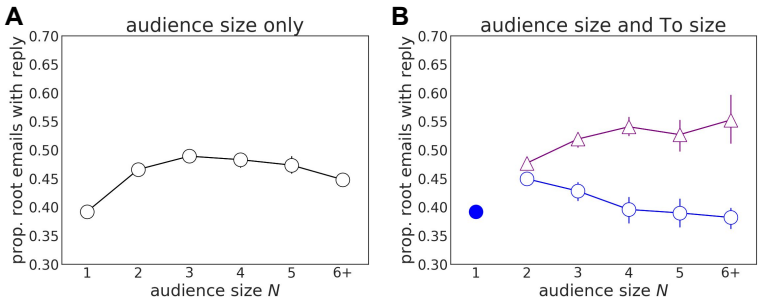

Fig. 4. Proportion of root emails receiving at least one reply. (A) Reply rate for various audience sizes N, replicating findings from Yang et al. [87]; (B) Reply rate for various audience configurations, accounting for audience size N and To size k: dyadic (●), N-flat (○), and N, 1-tiered (△).

4.5 Reply rate

Finally, we examine how an email's configuration relates to the subsequent interaction. Here, we focus on an email's propensity to receive replies—a simple but fundamental attribute of the interaction. Prior studies of email have developed models that predict replies as a practical task, but have given less consideration to the potential significance of an email's audience configuration. For instance, Kooti et al. [44] only analyzes dyadic emails; Yang et al. [87], which also examines the Avocado collection, accounts for audience size but not how the audience is organized and finds that reply rate increases up to $N = 3$ before decreasing, as depicted in Figure 4A. A finer-grained account of audience configurations could enable more intricate studies of interactions within groups, while pointing to additional features these predictive models could use.

Figure 4B shows the proportion of emails which receive at least one reply, across different configurations, such that both audience size N and To size k are taken into account. As with prior work, we note that the reply rate is lowest among dyadic emails; one possible explanation is that dyadic communications are relatively easy to address in another medium like telephone or face-to-face conversations [37, 57], so we do not observe responses via email.

We find a noticeable divergence based on an audience's organization: tiered emails have a higher reply rate than flat ones, and this gap widens as the audience size increases. Importantly, these contrasting trends are conflated in Figure 4A, which does not account for To size: the inflection point at $N = 3$ could reflect that for larger N, flat emails decrease in reply rate while increasing in frequency relative to tiered emails (as shown in Figure 1B). We can also disambiguate between different addressees: among tiered emails that receive replies, 71% versus 30% get replies from recipients in To and Cc respectively, indicating that such designations are closely tied to more active or passive roles in the communication.

Multiple interpretations are suggested by these findings. First, the relation between configuration and reply rate could reflect the situations involved; highly collaborative projects that spur active email exchanges may also involve multiple participants playing different roles. Second, a sender might tailor configurations to reflect particular intentions (as shown in Section 4.3) or to encourage some downstream consequence; including overseers could put pressure on the target of an email to respond. Third, these differences might also result from how recipients perceive a configuration. For instance, as N increases, the N recipients in a flat email may be less inclined to reply to the email, echoing the phenomenon of diffusion of responsibility [47]. This effect may be mitigated when the sender explicitly designates a primary target in a tiered configuration.

4.6 Sender-controlled protocol

We would like to tease apart two explanations for the patterns we have observed: the configurations, or the senders. For instance, configuration A may be more likely than configuration B to contain a request because senders in Avocado make meaningful and systematic distinctions between them; alternatively, different senders—with different propensities to make requests—also tend to choose different configurations. In order to gauge the extent to which our findings reflect meaningful differences in configuration choices, beyond sender idiosyncracies, we compare configurations in a sender-controlled fashion. We aggregate each attribute over each sender and configuration (e.g., for each sender, we compute the proportion of their dyadic emails with requests), and then compare attributes between configurations *paired on the sender*.[9]

Statistical significance under these paired comparisons is indicated in Table 2 using asterisks (*). Many differences still hold under this tighter protocol, suggesting that differences in *configuration* are meaningful. However, we also note some differences in the uncontrolled setting that aren't observed after pairing, suggesting that they might be contingent on the preferences of a limited subset of senders. For instance, N, 1-tiered emails are no more likely to have requests than 2-tiered emails (last column); a potential interpretation is that some senders who tend to issue requests, like managers of larger teams, also tend to address larger sets of recipients.

5 ANALYSIS: RECONFIGURATIONS

Thus far, we have examined how senders make choices about the inclusion and organization of recipients at the start of an email thread. However, as an interaction proceeds, people might modify the audience when they reply to a preceding message, by adding or removing participants. Here, we extend our analyses of email interactions to examine audience *reconfigurations*. Extending past accounts, which have drawn on close readings of small samples of reconfiguring emails [76] or on testimony from email senders [32], our approach surfaces characteristics that systematically recur in different types of reconfigurations across a larger corpus.

As a starting point, we focus on reconfigurations where the sender adds or removes recipients that were present in a preceding message; 35% of email replies involve such an action. We denote the number of added and removed recipients as $k_{\texttt{add}}$ and $k_{\texttt{rm}}$ respectively. As such, we abstract away from the address field from which previous recipients are removed, or to which new recipients are added; we also ignore cases where a sender moves recipients between `To` and `Cc`.[10]

We note that the possible reconfigurations that can occur are contingent on the configurations of preceding emails; in the extreme, a sender cannot remove more recipients than were initially present. To simplify our present analyses, we focus on the 16,136 reconfiguring emails which are replies to the *root* email of a thread, comprising 52% of all emails with a reconfiguration.[11]

5.1 Distribution of reconfigurations

We first examine the distribution of reconfigurations, in terms of $k_{\texttt{add}}$ and $k_{\texttt{rm}}$. Prior work has highlighted particular types of reconfigurations—replying to one person versus replying-all [32], or adding new recipients [76]. Here, we take a broader view, and consider the prevalence of these particular actions in relation to other types of reconfiguration. To structure our analysis, we divide

[9]We only consider sender-configuration pairs where the sender sent at least three emails of that configuration. Modifying this constraint yields qualitatively similar albeit somewhat noisier results.

[10]While future work could examine moving recipients between address fields, we note that it is difficult to distinguish between a sender's deliberate choice and automation from the user interface; many email clients automatically move recipients between `To` and `Cc` (e.g., when replying-all, all recipients except the previous sender may be moved to `Cc`).

[11]For comparison, 46% of replies overall are directed at the root of a thread, suggesting that reconfigurations are more likely to occur as a reply to a root email, rather than further down in a thread.

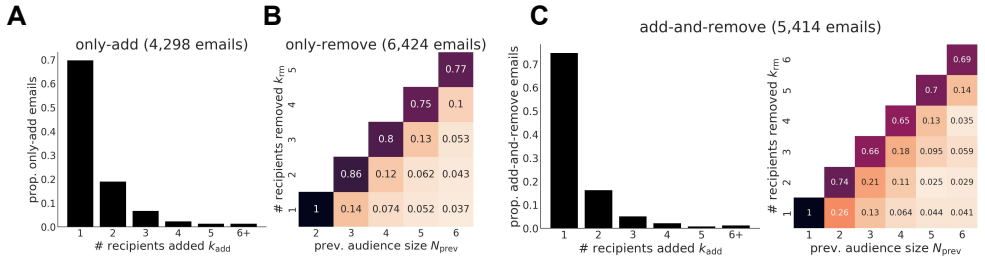

Fig. 5. Distribution of the number of added and removed recipients, $k_{\texttt{add}}$ and $k_{\texttt{rm}}$, in reconfiguring emails that are replies to the roots of threads. (A) distribution of $k_{\texttt{add}}$ for only-add emails (27% of reconfiguring emails); (B) distribution of $k_{\texttt{rm}}$ per previous audience size $N_{\texttt{prev}}$ for only-remove emails (40% of reconfiguring emails); (C) distribution of $k_{\texttt{add}}$ (left) and distribution of $k_{\texttt{rm}}$ per $N_{\texttt{prev}}$ (right) for add-and-remove emails (33% of reconfiguring emails).

reconfigurations into three cases: 1) the sender only adds recipients (4,298 or 27% of reconfiguring emails); 2) the sender only removes recipients (6,424 or 40%); 3) the sender both adds and removes recipients (5,414 or 33%).

Figure 5 shows the distributions of $k_{\texttt{add}}$ and $k_{\texttt{rm}}$ for each case. In both the only-add and add-and-remove cases (Figures 5A and 5C, left), we see that the distribution of $k_{\texttt{add}}$ is heavily skewed towards adding only a few new recipients.

Note that senders cannot remove more recipients than were previously present in the thread. We denote the number of previous recipients as $N_{\texttt{prev}}$, and compute the distribution of $k_{\texttt{rm}}$ per $N_{\texttt{prev}}$, which we visualize as heatmaps. In the only-remove case (Figure 5B), senders tend to remove $N_{\texttt{prev}}$-1 recipients (88% of only-remove emails), while in the add-and-remove case (Figure 5C, right), the sender tends to remove all $N_{\texttt{prev}}$ previous recipients (85% of add-and-remove emails). These cases, shown as the dark diagonals of the heatmaps, correspond to replying-one or forwarding an email. Both of these actions are explicitly implemented by email clients, suggesting that the choice of reconfiguring an audience may be skewed by what the user interface makes expedient.

5.2 Characterizing reconfigurations

We now extend our preceding analyses of root configurations to examine senders' decisions to reconfigure their audience. For our present study, we focus on two prominent types of reconfiguration identified above: adding at least one recipient without any removals—henceforth `Add`'s; replying to one recipient while removing the others, without any additions—henceforth `Reply-1`'s. We contrast emails where a sender has reconfigured the preceding audience with *stable* emails where the participants are left unchanged.

Here, we seek to characterize a sender's decision to *modify* an existing configuration. However, naively comparing `Add`, `Reply-1` and stable emails would conflate factors underlying the modification with factors reflecting the present configurations of the emails being examined. Concretely, a `Reply-1` and `Add` email may differ because senders have different intentions in removing or adding recipients, or because they have different intentions in sending dyadic versus multi-recipient emails.

To circumvent this problem, we *pair* each reconfiguring email with a stable email that has the same audience size N and `To`-size k. We ensure that the stable reference point is also a reply to a root email, such that our findings do not conflate properties of a reconfiguration decision with dynamics of an email thread. For instance, this means that we compare each `Reply-1` to the second email of a dyadic exchange. Per Section 4.6, we enforce that both emails in a pair are from the same sender. This procedure results in 2,276 pairs for `Add`'s and 4,792 pairs for `Reply-1`'s.

	attribute	Add	Reply-1
(i)	has request	↑↑↑	
	has commitment		↓↓↓
(ii)	has recipient name	↑↑↑	↑↑↑
(iii)	wordcount	↑↑↑	↓↓
	% 1st-person sing.	↓↓↓	
	% 2nd-person	↑↑↑	
	% 1st-person pl.		
(iv)	has reply	↑↑↑	↑↑↑

Table 3. Significance test statistics on email content attributes and reply rate, comparing Add and Reply-1 emails with paired stable emails. Arrow direction indicates whether the reconfiguring emails are greater (↑) or smaller (↓) in that attribute than the stable emails. The number of arrows indicates Bonferroni-corrected significance under the Wilcoxon signed-rank test (real-valued attributes), or McNemar's test (binary attributes).

Table 3 shows comparisons between reconfiguring emails and their stable counterparts. We focus first on Add emails, where a sender includes a new recipient. Compared to stable emails, Add's more frequently contain requests and mentions of recipient names. These differences mirror contrasts we have previously observed between tiered and flat root emails; indeed, one interpretation is that new recipients are implicitly singled out when they are added to a thread, similar to designating a particular recipient in To (e.g., consider looping in a new additional participant so that they can fulfill a request). Prior work has suggested that senders may add recipients to facilitate their later participation [76]; we correspondingly observe that Add's are also more prone to receive replies.

Next, we examine Reply-1 emails. Our findings suggest two interpretations of this action. First, replying-one removes recipients that a sender deems to be less relevant to a message, perhaps to facilitate a more informal interaction with fewer overseers [32]; this may account for briefer messages with fewer commitments. Second, in replying-one, a sender also singles out the remaining participant, thus directing the focus of the message on them; this may be reflected in an increased propensity for calling out the recipient by name, or for this recipient to subsequently acknowledge the message by replying. As with configurations of root emails, we suggest that reconfigurations could be seen in part as a signaling mechanism (e.g., to denote a more targeted, actionable message).

5.3 Relation to initial configurations

A sender who initiates an email thread may not be the same person as a replier who subsequently reconfigures the audience. As such, accounting for audience configurations highlights a potential source of tension: do senders who initiate an email thread anticipate later reconfigurations? Prior work on reply-all practices suggest two possibilities [32]: Senders may intend for a subsequent reply-all since they deliberately included multiple recipients; they may likewise intend for a reply-one, for instance by explicitly specifying this intent in their email. On the other hand, a replier may ignore or misunderstand these intentions, leading to frustrations. This potential mismatch between sender and replier intentions pertains to subsequent Add's as well: a recipient who is added later on in the thread is able to view preceding emails which weren't originally addressed to them, which could be problematic if such a reconfiguration isn't sanctioned by the initial sender.

As noted in [32], very few studies have systematically examined such tensions, even as there is some popular understanding of them (e.g., consider the prevalent antipathy towards replying-all). As a starting point, we explore the relation between a sender and later replier's choice of audience:

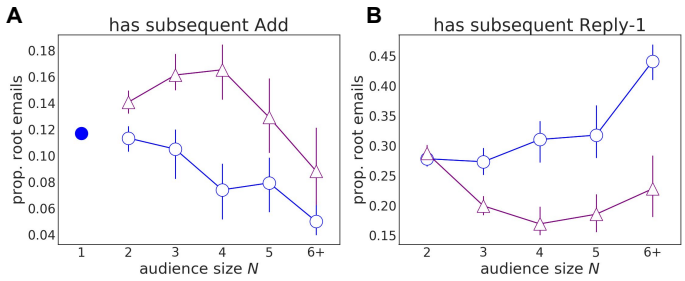

Fig. 6. Proportion of root emails whose replies include (A) an Add and (B) a Reply-1, for dyadic (•), N-flat (○), and N, 1-tiered (△) emails.

for each type of root configuration, we compute the proportion of emails that receive a reply with an Add or Reply-1.[12] Here, we focus on the 34,114 root emails with at least one reply.

Figures 6A and 6B show the proportion of root emails with different configurations, that later get a reply with Add or Reply-1 reconfigurations. We see that the propensity for subsequent Add's is higher for tiered than for flat roots (Fisher's test $p < 0.05$ at $N = 2$); recalling the intuitively task-oriented nature of tiered emails, perhaps repliers may be more prone to include additional stakeholders, e.g., to see through a collaborative effort or request.

The propensity for Reply-1's is highest for N-flat roots with large N, echoing the idea that senders use Reply-1's to pare away less relevant participants in such untargeted exchanges (Fisher's test $p < 0.001$, comparing 2-flat to N-flat emails). In contrast, this propensity is lower for N, 1-tiered roots: perhaps the more complex organization of recipients signals that the initial sender has deliberately included participants they deem to be relevant to the communication. In particular, this means that the recipients in Cc, while auxiliary in the sense that they might not be expected to actively participate in the thread, nonetheless tend to be retained in the ensuing interaction. These findings suggest that some email threads are more expected than others to involve audiences that are subject to change, opening up avenues for future work to more closely examine these expectations and cases where they might later be violated.

6 DISCUSSION

In this work, we examined how people within an organization configure their audience when communicating, via a case study of emailing behavior. In our setting, our analyses point to several systematic dependencies between the types of configurations chosen by email senders and other aspects of the communication. More broadly, we provided a general description of audience configurations as an important and multifacted component of social communication: configurations choices are salient in channels beyond email that likewise enable speakers to make explicit decisions about their audience. We see our present study as a starting point to frame future research on audience configurations; in what follows, we highlight some key takeaways as well as limitations that suggest directions for future work.

Throughout our analyses, we have demonstrated the utility of studying audience configuration decisions as a complement to examining the content of messages, which has been given more prominent treatment in past studies of communication behavior. In our setting, we find meaningful distinctions among audience configurations which underline that drawing finer comparisons

[12]Given our data filtering decisions, as detailed in Section 3.1, we ignore self-replies; as such, the senders of the reconfiguring emails considered in this analysis are different from the senders of the root emails.

between different types of configurations, rather than abstracting away such audience-centric choices, could enrich accounts of how people communicate on several fronts. Notably, we find that the propensity of emails to receive replies varies in striking ways with the configuration involved (Sections 4.5 and 5.2); such differences in reply propensity would have been overlooked without a granular account of configurations. We likewise find clear contrasts in the content of messages that are directed at different configurations (Section 4.4) or that accompany different reconfigurations (Section 5.2). Indeed, as our exploration of sender-recipient relationships (Section 4.2) suggests, senders are directly faced with configuration decisions in situations such as contacting multiple individuals across an organization; our framing brings these decisions to the fore of the analyses.

We have identified a few key dimensions along which audiences could be configured. We highlight the audience's organization into different recipient roles, recalling theoretical accounts of the distinction between primary addressees and auxiliary participants [17, 31]. To clarify and extend these accounts, we note an interesting subtlety in interpreting the importance of auxiliary audience members. We show that they tend not to be explicitly called out in an email, and their placement in `Cc` perhaps signals that the sender does not expect them to actively participate in a reply. However, this seemingly passive role does not make them superfluous: tiered emails where the sender has decided to include them exhibit contrasts on several fronts with flat emails where no recipient is designated as auxiliary in the address line. We have also highlighted salient patterns tied to a sender's choice to reconfigure the audience in a reply. In accounting for the potentially dynamic nature of the audience in an interaction, we point to an interesting and understudied interplay between senders' and repliers' intentions (Section 5.3) that future work could further investigate, perhaps in the vein of communicative misunderstandings [12, 16, 80].

In framing audience configurations as decisions that people can explicitly make, we offer additional interpretations of the relationship between a message and its audience that go beyond the audience design literature, which has generally focused on how people design messages in reaction to a given audience. The audience and message may reinforce one another, as was observed in how email senders mention recipient names (Sections 4.4 and 5.2). The audience's configuration may also supplement information conveyed in the language of the message, for instance in including relevant participants and designating them as auxiliaries. As such, a configuration may be seen as a complementary source of signal to cues found in the message itself [61, 77]. In conjunction with the performative nature of messages, speakers could use configurations to precipitate an outcome: for instance, they could encourage a reply by highlighting primary recipients in tiered configurations (Section 4.5) or modifying the audience to target a single person (Section 5.2), hence removing uncertainty about the focus of attention and the designation of responsibility.

More intricate studies of the message and the audience could clarify the nature of the link between them. For instance, our analysis of recipient names could be extended to account for the rhetorical functions of different name-mentions, hence clarifying how speakers designate the roles of their various addressees. We could also investigate potential causal effects of choosing different configurations to more rigorously understand the utility of configuration choices—for instance, experimental studies could vary the audience configuration of a fixed message and observing the impact of such interventions on the subsequent interaction.

We note that our present analyses miss an important aspect of configurations: the precise identities of audience members. Future work could address this gap, perhaps by using similar abstractions of roles and relationships as in our exploration of company divisions.

Analytic implications. We supplement these broad suggestions by describing some concrete ways to extend existing studies of communication behavior by accounting for audience configurations. First, we see a range of other analyses that could be augmented by disambiguating between configurations that people have chosen. For instance, records of communications have been used

to induce networks that represent organizational and social structures, and that can then shed light on phenomena such as power relations [54] or responses to events [5, 24, 70]. To facilitate a large-scale examination of communication networks, these studies have generally abstracted away from the particular characteristics of different interactions, coarsely drawing links between people whenever any communication occurs between them. Our findings show that the relationships between participants in an interaction are reflective of the configuration that recipients of a message are organized into, suggesting that accounting for configurations could enrich representations of communications networks. For instance, edges between people in a network of email exchanges could disambiguate between the different communicative roles that recipients tend to be placed in by the senders, hence encoding designations that reflect the underlying heterogeneity of the relationships involved.

Second, we see promise in exploring the configuration of audiences via the lens of pragmatics and social psychology. For instance, past work has examined how the success of requests could be tied to linguistic attributes that signal politeness or interpersonal affinity [3, 86]. Our analyses have suggested that email senders configure the group of recipients of a request in particular ways. A future study could take up these observations to investigate the relation between audience-centric attributes (such as the presence of an overseer in `Cc`) and the success of a request, enabling analyses that more holistically account for the various types of decisions people make when issuing requests, perhaps to increase the chances of fulfillment. It would also be fruitful to substantiate many other intuitions about the psychological interpretation of audience configurations, such as avoiding reply-all's as a courtesy [32], or perceiving the inclusion of supervisors in `Cc` as an act of control [35]. Here, we could draw on approaches to developing linguistic characterizations of social behaviors such as politeness [9, 22], power dynamics [28, 61, 66], emotional valence [12], formality [64] and acculturation [25, 78].

Design implications. Accounting for audience configurations could also inform various proposals for designing better communications systems. Past studies have suggested various ways to streamline the task of emailing [62], for instance by ranking the importance of messages [39] or extracting intents expressed within messages [13, 71, 82]. Audience configurations could be explored as an additional source of signal for tools that automate these tasks. The configuration of an audience could also be highlighted to users to guide them in managing their inbox: past work has suggested that being included in the `Cc` line is a signal of lower priority [39, 72], but other signals—such as being a recipient in `To` with several others in `Cc`, or being the remaining participant in a message with a `Reply-1`—could be informative as well. Such applications are also potentially applicable in other settings where the audience is configurable, such as group chat.

Understanding how people might use audience configurations could also guide the design of new affordances which provide speakers some ability to tailor their audience. We have highlighted various dimensions of configurations that other platforms could potentially support; for instance, a platform might consider enabling people to specify different tiers of addressees, replicating the particularly explicit distinction between `To` and `Cc` in email. Our analyses suggest various implications of supporting tiered audience configurations; such a feature might enable message-senders to convey additional signal and set expectations for subsequent replies. Additional studies would be needed to gain a better sense of whether such affordances would be appropriate for a particular platform. For instance, while the organizational structure of a company might justify the arrangement of addressees into different communicative tiers, such a decision might be more fraught in settings without explicitly-stated subdivisions and hierarchies.

6.1 Limitations and future work

Our study has several limitations that suggest other opportunities for future work. First, we have assumed throughout that the configuration of the audience of email messages reflects an explicit and deliberate communicative choice. We have argued that this assumption is valid, relative to other communication channels where such choices are more constrained. However, other factors beyond the sender's communicative considerations may shape how configurations are used. As we have seen, the distribution of reconfigurations is strongly tied to options exposed by email clients (Section 5.1); automated addressee recommendations as presently implemented by Gmail and Outlook could further shape sender behaviors. A sender's choices are also subject to various norms, especially in routinized settings such as companies in which actions such as replying-all [32] or including certain recipients [73] might be encouraged or disfavored. Future work could provide characterizations of cultural or organizational norms that encompass audience configurations, and more closely investigate their role in a sender's decision.

While our study focuses on the basic `To` and `Cc` mechanisms, email provides functionality to send messages to less explicitly specified audiences as well: senders can target mailing lists, or put recipients in the `Bcc` field. Extending our present analyses to understand when these options are used could add nuance to our formulation of audience-centric decisions, allowing us to draw connections to theoretical accounts of other phenomena such as overhearers [18] or to settings involving less explicit audiences, such as large group chat channels [36] or online discussions [88].

Our computational approach limits us to high-level post-hoc analyses of configuration choices and inevitably misses the nuances of particular situations, especially for the less frequent configuration types that we have largely ignored. We also have a limited understanding of senders' motivations and recipients' perceptions of a configuration. To fill this gap, we could build on the interview and discourse analysis-based studies that have informed our present study, and examine a wider range of audience-related considerations.

While we have argued that our corpus is an illustrative case example, it represents the emailing practices of one company, and we expect that people in different contexts could use configurations differently. We have already seen one example of this potential variability: in Section 4.2, we observe qualitatively similar trends for emails originating in sales and engineering, but the magnitudes of these effects diverge, perhaps due to the different roles, organizational structures and norms within these subdivisions. Future work could address such contextual factors and apply our approach to analyze other corpora, to better understand the variation across different settings.

Configuring the audience is one of many concurrent decisions that people make when communicating. For instance, people can also specify the communication medium [37], moving between emailing, instant-messaging, phone calls, walking over to someone's desk, and calling a meeting [57]. Thus, there is ample opportunity for future work to examine audience configurations in the broader ecosystem of communication and collaboration.

ACKNOWLEDGMENTS

The authors thank the members of the IDEAS and ASI groups at Microsoft Research for the many insightful discussions, and Ali Alkhatib, Jonathan P. Chang, Caleb Chiam, Dan Cosley, Cristian Danescu-Niculescu-Mizil, Liye Fu, Ahmed Hassan Adwallah, Jack Hessel, Malte Jung, Gilly Leshed, Samir Passi, and the anonymous reviewers for their extremely helpful comments. The first author was supported by a Microsoft Research PhD Fellowship.

REFERENCES

[1] 2019. *Email Statistics Report, 2019-2023.* Technical Report. Radicati Group.
[2] Sakhar Alkhereyf and Owen Rambow. 2017. Work Hard, Play Hard: Email Classification on the Avocado and Enron Corpora. In *Proceedings of TextGraphs*.
[3] Tim Althoff, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2014. How to Ask for a Favor: A Case Study on the Success of Altruistic Requests. In *Proceedings of ICWSM*.
[4] John L. Austin. 1975. *How to Do Things with Words*.
[5] Brandy L. Aven. 2015. The Paradox of Corrupt Networks: An Analysis of Organizational Crime at Enron. *Organization Science* (2015).
[6] Natalya N. Bazarova and Yoon Hyung Choi. 2014. Self-Disclosure in Social Media: Extending the Functional Approach to Disclosure Motivations and Characteristics on Social Network Sites. *Journal of Communication* (2014).
[7] Allan Bell. 1984. Language Style as Audience Design. *Language in Society* (1984).
[8] Michael S. Bernstein, Eytan Bakshy, Moira Burke, and Brian Karrer. 2013. Quantifying the Invisible Audience in Social Networks. In *Proceedings of CHI*.
[9] Penelope Brown and Stephen C Levinson. 1987. *Politeness: Some Universals in Language Usage*.
[10] Anne Marie Bülow, Joyce Y. H. Lee, and Niki Panteli. 2019. Distant Relations: The Affordances of Email in Interorganizational Conflict. *International Journal of Business Communication* (2019).
[11] Ronald S. Burt. 2004. Structural Holes and Good Ideas. *Amer. J. Sociology* (2004).
[12] Kristin Byron. 2008. Carrying Too Heavy a Load? The Communication and Miscommunication of Emotion by Email. *The Academy of Management Review* (2008).
[13] Vitor Carvalho and William Cohen. 2006. Improving "Email Speech Acts" Analysis via N-Gram Selection. In *Proceedings of the ACTS Workshop*.
[14] Vitor R. Carvalho. 2011. Modeling Intention in Email - Speech Acts, Information Leaks and Recommendation Models. In *Studies in Computational Intelligence*.
[15] Justine Cassell and Dona Tversky. 2005. The Language of Online Intercultural Community Formation. *Journal of Computer-Mediated Communication* (2005).
[16] Jonathan P Chang, Justin Cheng, and Cristian Danescu-Niculescu-Mizil. 2020. Don't Let Me Be Misunderstood:Comparing Intentions and Perceptions in Online Discussions. In *Proceedings of WWW*.
[17] Herbert H Clark and Thomas B Carlson. 1982. Hearers and Speech Acts. *Language* (1982).
[18] Herbert H. Clark and E. Schaefer. 1987. Concealing One's Meaning from Overhearers. *Journal of Memory and Language* (1987).
[19] William W. Cohen, Vitor R. Carvalho, and Tom M. Mitchell. 2004. Learning to Classify Email into "Speech Acts". In *Proceedings of EMNLP*.
[20] James S. Coleman. 1988. Social Capital in the Creation of Human Capital. *Amer. J. Sociology* (1988).
[21] Simon Corston-Oliver, Eric Ringger, Michael Gamon, and Richard Campbell. 2004. Integration of Email and Task Lists. (2004).
[22] Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. A Computational Approach to Politeness with Application to Social Factors. In *Proceedings of ACL*.
[23] Michal Danilák. 2020. Langdetect. https://github.com/Mimino666/langdetect.
[24] Jana Diesner, Terrill L. Frantz, and Kathleen M. Carley. 2005. Communication Networks from the Enron Email Corpus "It's Always About the People. Enron Is No Different". *Computational and Mathematical Organization Theory* (2005).
[25] Gabriel Doyle, Amir Goldberg, Sameer Srivastava, and Michael Frank. 2017. Alignment at Work: Using Language to Distinguish the Internalization and Self-Regulation Components of Cultural Fit in Organizations. In *Proceedings of ACL*.
[26] Nicolas Fay, Simon Garrod, and Jean Carletta. 2000. Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size. *Psychological Science* (2000).
[27] Casey Fiesler, Munmun De Choudhury, Eric Gilbert, Michaelanne Dye, Jessica L. Feuston, Chaya Hiruncharoenvate, C.J. Hutto, Shannon Morrison, Parisa Khanipour Roshan, Umashanthi Pavalanathan, and Amy S. Bruckman. 2017. What (or Who) Is Public?: Privacy Settings and Social Media Content Sharing. In *Proceedings of CSCW*.
[28] Eric Gilbert. 2012. Phrases That Signal Workplace Hierarchy. In *Proceedings of CSCW*.
[29] Eric Gilbert and Karrie Karahalios. 2009. Predicting Tie Strength with Social Media. In *Proceedings of CHI*.
[30] Erving Goffman. 1959. *The Presentation of Self in Everyday Life*.
[31] Erving Goffman. 1979. Footing. *Semiotica* (1979).
[32] Sukeshini A. Grandhi and Lyndsey Lanagan-Leitzel. 2016. To Reply or To Reply All: Understanding Replying Behavior in Group Email Communication. In *Proceedings of CSCW*.
[33] Mark S. Granovetter. 1973. The Strength of Weak Ties. *Amer. J. Sociology* (1973).

[34] Catherine Grevet, David Choi, Debra Kumar, and Eric Gilbert. 2014. Overload Is Overloaded: Email in the Age of Gmail. In *Proceedings of CHI*.
[35] Tessa Haesevoets, David De Cremer, Leander De Schutter, Jack McGuire, Yu Yang, Xie Jian, and Alain Van Hiel. 2019. Transparency and Control in Email Communication: The More the Supervisor Is Put in Cc the Less Trust Is Felt. *Journal of Business Ethics* (2019).
[36] Mark Handel and James D Herbsleb. 2002. What Is Chat Doing in the Workplace?. In *Proceedings of CSCW*.
[37] Caroline Haythornthwaite. 2001. Exploring Multiplexity: Social Network Structures in a Computer-Supported Distance Learning Class. *The Information Society* (2001).
[38] Matthew Honnibal and Ines Montani. 2020. spaCy: Natural Language Understanding with Bloom Embeddings, Convolutional Neural Networks and Incremental Parsing.
[39] Eric J. Horvitz, Andy Jacobs, and David Hovel. 1999. Attention-Sensitive Alerting. In *Proceedings of UAI*.
[40] Paula Jarzabkowski and David Seidl. 2008. The Role of Meetings in the Social Practice of Strategy. *Organization Studies* (2008).
[41] Ian Johnson. 2013. Audience Design And Communication Accommodation Theory: Use Of Twitter By Welsh–English Biliterates. In *Social Media and Minority Languages: Convergence and the Creative Industries*.
[42] Ewa Kacewicz, James W. Pennebaker, Matthew Davis, Moongee Jeon, Arthur C, Ewa Kacewicz, James W. Pennebaker, Matthew Davis, Moongee Jeon, and Arthur C. Graesser. 2013. Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology* (2013).
[43] Bryan Klimt and Yiming Yang. 2004. Introducing the Enron Corpus. In *CEAS*.
[44] Farshad Kooti, Luca Maria Aiello, Mihajlo Grbovic, Kristina Lerman, and Amin Mantrach. 2015. Evolution of Conversations in the Age of Email Overload. In *Proceedings of WWW*.
[45] Gueorgi Kossinets and Duncan J. Watts. 2009. Origins of Homophily in an Evolving Social Network. *Amer. J. Sociology* (2009).
[46] Robert E Kraut and Paul Attewell. 1996. Media Use in a Global Corporation: Electronic Mail and Organizational Knowledge. *Culture of the Internet* (1996).
[47] Bibb Latané. 1981. The Psychology of Social Impact. *American Psychologist* (1981).
[48] Yao Li, Xinning Gui, Yunan Chen, Heng Xu, and Alfred Kobsa. 2018. When SNS Privacy Settings Become Granular: Investigating Users' Choices, Rationales, and Influences on Their Social Experience. In *Proceedings of CSCW*.
[49] Eden Litt. 2012. Knock, Knock. Who's There? The Imagined Audience. *Journal of Broadcasting & Electronic Media* (2012).
[50] Yabing Liu and Krishna P Gummadi. 2011. Analyzing Facebook Privacy Settings: User Expectations vs. Reality. In *Proceedings of IMC*.
[51] Xiao Ma, Jeff Hancock, and Mor Naaman. 2016. Anonymity, Intimacy and Self-Disclosure in Social Media. In *Proceedings of CHI*.
[52] Ifigeneia Machili, Jo Angouri, and Nigel Harwood. 2019. 'The Snowball of Emails We Deal With': CCing in Multinational Companies. *Business and Professional Communication Quarterly* (2019).
[53] Alice E. Marwick and danah boyd. 2010. I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society* (2010).
[54] Andrew McCallum, Xuerui Wang, and Andrés Corrada-Emmanuel. 2007. Topic and Role Discovery in Social Networks with Experiments on Enron and Academic Email. *Journal of Artificial Intelligence Research* (2007).
[55] Tanushree Mitra and Eric Gilbert. 2013. "Analyzing Gossip in Workplace Email" by Tanushree Mitra and Eric Gilbert, with Ching-Man Au Yeung as Coordinator. *ACM SIGWEB Newsletter* (2013).
[56] Sean A. Munson, Karina Kervin, and Lionel P. Robert, Jr. 2014. Monitoring Email to Indicate Project Team Performance and Mutual Attraction. In *Proceedings of CSCW*. New York, NY, USA.
[57] Bonnie A. Nardi, Steve Whittaker, and Erin Bradner. 2000. Interaction and Outeraction: Instant Messaging in Action. In *Proceedings of CSCW*.
[58] Dong Nguyen, Dolf Trieschnigg, and Leonie Cornips. 2015. Audience and the Use of Minority Languages on Twitter. In *Proceedings of ICWSM*.
[59] Doug Oard, William Webber, David Kirsch, and Sergey Golitsynskiy. 2015. Avocado Research Email Collection.
[60] Jacob Palme. 1997. Common Internet Message Headers. https://tools.ietf.org/html/rfc2076.
[61] Niki Panteli. 2002. Richness, Power Cues and Email Text. *Information & Management* (2002).
[62] Soya Park, Amy X. Zhang, Luke S. Murray, and David R. Karger. 2019. Opportunities for Automating Email Processing: A Need-Finding Study. In *Proceedings of CHI*.
[63] Umashanthi Pavalanathan and Jacob Eisenstein. 2015. Audience-Modulated Variation in Online Social Media. *American Speech* (2015).
[64] Ellie Pavlick and Joel Tetreault. 2016. An Empirical Analysis of Formality in Online Communication. *TACL* (2016).

[65] James W. Pennebaker, Cindy K. Chung, Joey Frazee, Gary M. Lavergne, and David I. Beaver. 2014. When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLOS ONE* (2014).
[66] Vinodkumar Prabhakaran and Owen Rambow. 2013. Written Dialog and Social Power: Manifestations of Different Types of Power in Dialog Behavior. In *Proceedings of IJCNLP*.
[67] Kristen Purcell. 2011. Search and Email Still Top the List of Most Popular Online Activities.
[68] Mark T. Rivera, Sara B. Soderstrom, and Brian Uzzi. 2010. Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms. *Annual Review of Sociology* (2010).
[69] Shane L. Rogers, Nicolas Fay, and Murray Maybery. 2013. Audience Design through Social Interaction during Group Discussion. *PLOS ONE* (2013).
[70] Daniel M. Romero, Brian Uzzi, and Jon Kleinberg. 2016. Social Networks Under Stress. In *Proceedings of WWW*.
[71] M. Sappelli, G. Pasi, S. Verberne, M. de Boer, and W. Kraaij. 2016. Assessing E-Mail Intent and Tasks in e-Mail Messages. *Information Sciences* (2016).
[72] Bahareh Sarrafzadeh, Ahmed Hassan Awadallah, Christopher H. Lin, Chia-Jung Lee, Milad Shokouhi, and Susan T. Dumais. 2019. Characterizing and Predicting Email Deferral Behavior. In *Proceedings of WSDM*.
[73] Joseph Schmitz and Janet Fulk. 1991. Organizational Colleagues, Media Richness, and Electronic Mail: A Test of the Social Influence Model of Technology Use. *Communication Research* (1991).
[74] John R. Searle. 1976. A Classification of Illocutionary Acts. *Language in Society* (1976).
[75] Georg Simmel. 1902. The Number of Members as Determining the Sociological Form of the Group. *Amer. J. Sociology* (1902).
[76] Karianne Skovholt and Jan Svennevig. 2006. Email Copies in Workplace Interaction. *Journal of Computer-Mediated Communication* (2006).
[77] Lee Sproull and Sara Kiesler. 1986. Reducing Social Context Cues: Electronic Mail in Organizational Communication. *Management Science* (1986).
[78] Sameer B. Srivastava, Amir Goldberg, V. Govind Manian, and Christopher Potts. 2018. Enculturation Trajectories: Language, Cultural Adaptation, and Individual Outcomes in Organizations. *Management Science* (2018).
[79] Yla R. Tausczik and James W. Pennebaker. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology* (2010).
[80] Jenny Thomas. 1983. Cross-Cultural Pragmatic Failure. *Applied Linguistics* (1983).
[81] Fernanda B. Viégas. 2005. Bloggers' Expectations of Privacy and Accountability: An Initial Survey. *Journal of Computer-Mediated Communication* (2005).
[82] Wei Wang, Saghar Hosseini, Ahmed Hassan Awadallah, Paul N. Bennett, and Chris Quirk. 2019. Context-Aware Intent Identification in Email Conversations. In *Proceedings of SIGIR*.
[83] Steve Whittaker, Victoria Bellotti, and Jacek Gwizdka. 2006. Email and PIM: Problems and Possibilities. *Communications in the ACM* (2006).
[84] Steve Whittaker and Candace Sidner. 1996. Email Overload: Exploring Personal Information Management of Email.. In *Proceedings of CHI*.
[85] Shaomei Wu, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. 2011. Who Says What to Whom on Twitter. In *Proceedings of WWW*.
[86] Diyi Yang and Robert E. Kraut. 2017. Persuading Teammates to Give: Systematic versus Heuristic Cues for Soliciting Loans. In *Proceedings of CSCW*.
[87] Liu Yang, Susan T. Dumais, Paul N. Bennett, and Ahmed Hassan Awadallah. 2017. Characterizing and Predicting Enterprise Email Reply Behavior. In *Proceedings of SIGIR*.
[88] Justine Zhang, Cristian Danescu-Niculescu-Mizil, Christina Sauper, and Sean J. Taylor. 2018. Characterizing Online Public Discussions Through Patterns of Participant Interactions. In *Proceedings of CSCW*.

A SUPPLEMENTARY MATERIAL

Here, we provide further details on the nature of our dataset of emails and the steps we took to process it, expanding on our description in Section 3.1. Code that implements this procedure can be found at https://github.com/tisjune/avocado-data-processing.

A.1 Data description

The Avocado email collection is comprised of data taken from 279 Microsoft Outlook accounts belonging to employees of a company referred to as Avocado. These records comprise the complete set of emails that were in the accounts at the time that the data was collected, along with some other artifacts like contact lists and calendar entries.

In the original LDC data release [59], the data was structured as XML files, with each entry corresponding to an email that was present in some folder in an employee's account. Thus, one could naively use the collection by simply using each XML entry as an email. This presented two challenges which our postprocessing steps sought to circumvent.

First, if employees deleted emails, then these would not be present in the collection as explicit XML entries. In particular, this meant that many emails lacked information about the preceding thread, since the email they were replying to was deleted. Second, there were many duplicate emails, since copies of an email could exist in the accounts of each of its senders and recipients.

A.2 Data processing

Our data processing procedure deals with each of these challenges in turn. Here, we refer to the emails surfaced by the XML entries as "original" emails and new emails recovered by our process as "inferred" emails. Inspecting subsamples of the data suggests that these steps produced a more complete set of emails in terms of email count and reply structure, and successfully removed the majority of duplicates.

Recovering deleted emails and thread structure. To augment the dataset with deleted emails and missing reply information, we made use of the full text of the original emails. For replies, this text often contained both the text of the reply, and the preceding emails separated by "Original Message" blocks. Thus, even though these preceding emails were deleted, we could still recover the thread via text processing.

We separated the text associated with an original email into blocks delimited by "Original Message", and for each inferred email took the parent to which it was replying to be the inferred email in the "Original Message" block directly underneath it. For each email, we extracted the subject line, recipients, timestamp, and body; we also parsed out prefixes on the subject line like "RE:" and "FWD:" using a regular expression.

Deduplicating emails. We deduplicated the emails in two passes. In the first pass, we used the email subject (with prefixes like "RE:" removed), sender address, recipient addresses, and timestamp as a key on which to deduplicate emails.

One complication is that emails are sometimes from different timestamps, which aren't always explicitly listed in the text. As such, identical emails could have two different timestamps, depending on whose inbox it came from (e.g., if people were in different time zones). To deal with this, we performed a second round of deduplication: we reconstructed the thread structure of the email, and for each set of replies to the same parent, deduplicated on just the subject line and sender/recipients.

Caveats on data completeness. If an email was deleted from all accounts involved, and did not have any replies, then we would not be able to recover it in our collection. As such, while we claim that the collection of emails we obtained after our postprocessing procedure is fairly thorough, and our analysis is representative of a significant portion of emailing activity, we note that if certain types of emails are systematically deleted then we would miss these types.

We also note that if recipients were `Bcc`'d, then their names would not show up in inferred emails (since by definition, `Bcc` hides these addressees from the other recipients). As such, we omitted all emails with any trace of `Bcc`-ing activity from our analyses (i.e., if an original email had a `Bcc`, we would ignore all the emails in the same thread).

Finally, some emails which are replies did not include the text of the original message below it. We did not consider an email to be the root of a thread if its subject line had a prefix like "RE:" in it, even if it did not have a parent email in the collection.

Data statistics. The original collection had 614K non-duplicate emails (and 938K before deduplication). Our augmented collection has 702K emails after deduplication (1,800K before). Our most

significant increase is in the reply structure. The original collection contained 121K emails which were replies to a preceding email; our collection contains 297K emails which are replies.

Subsetting decisions. As described in Section 3.1, our analyses were restricted to a subset of the data. Here, we provide more details about the various subsetting decisions we made, and the number of emails that were omitted as a result. Note that an email could be omitted for multiple reasons, such that the percentages reported below refer to potentially-overlapping subsets of emails.

- We considered emails sent only among Avocado employees (identified via the procedure detailed in the next section). While other work has fruitfully considered communications with people outside of organizations [70], we omitted such emails to simplify our present analyses; in addition, it was difficult to distinguish between external senders and recipients who were actual people, versus mailing lists or automated services. This step removed 65% of emails in 70% of threads. Manual inspection suggests that many of these omitted messages come from automated senders (such as newsletters or promotions). In other words, we argue that these omissions, while numerically substantial, consist in large part of emails which we would not interpret as social communications between people.
- We ignored emails sent to mailing lists within Avocado, and emails involving non-human senders (such as automatically-generated bug reports). In addition to facilitating our initial analyses, this decision reflects two limitations: it was sometimes difficult to distinguish between these two types of addresses, and we were unable to recover the identities of people on each mailing list at the time that each email was sent. This step removed 13.4% of emails in 9.91% of threads.
- We ignored self-emails where the sender is the only recipient, thus omitting 7.7% of emails.
- As noted above, we ignored emails in threads where at least one message was addressed to people in the `Bcc` field. Approximately 0.5% of emails were omitted as a result of this filter, though as we have described, there are probably more emails with `Bcc`'s in the data.

A.3 Additional postprocessing steps

We also took some additional steps to prepare the data for our analyses in Sections 4 and 5.

Identifying employee information. We took steps to extract a more complete set of employees (from the 279 accounts originally in the collection) and match them to information about their email addresses, names and nicknames, and roles.

In particular, we extracted all (name, address) pairs from emails where the address had an "@avocadoit" substring; we kept a pair if it occurred at least 5 times. Since multiple names and multiple addresses could be associated with each other, we derived sets of names and addresses corresponding to each employee by finding all the connected components of the graph induced by edges from a name to an address in each pair. We then manually curated this list to remove inaccurate pairs that resulted in multiple people getting mapped to one component, and merged some components if they corresponded to the same person. We manually identified @avocadoit entities corresponding to system accounts (e.g., conference rooms that could be booked) or mailing lists, rather than employees. This process resulted in 320 employees being identified.

In order to perform our analyses of name mentions, we augmented the original set of names associated with each person with commonly-used nicknames. We computed the pointwise mutual information (PMI) between each employee and each word in the emails they received, restricting to words which started with the same letter as the recipient's original name. We considered all (word, name) pairs above a minimum frequency and PMI; we then manually filtered these to arrive at a set of nicknames.

To find job titles, we used information from contacts lists included with the collection. In a minority of cases, where these titles were ambiguous, we made use of text in the signatures that an employee appended to an email (e.g., "John Smith, Account Manager"). Note that job titles are subject to change; for our analyses, we assume that the *subdivision* that the employee is associated with is fairly stable. With additional effort, it might be possible to extract and examine information related to status in the organization, though the data doesn't contain any org charts, and we expect power relations to change as people moved around the company.

We removed emails involving mailing lists from our present analyses; a further augmentation we did not pursue could be to infer the identities of the people on these lists.

Processing address fields. In order to analyze the configuration of address fields, we postprocessed the address fields of the emails by deduplicating addressees. First, we removed the email sender from all of the recipient fields; next, we removed all `Cc`'d recipients who were also in the `To` field; next, we deduplicated addressees within each address field (i.e., if the email was sent to more than one of their addresses). If all recipients were in `Cc`, we treated them as all being in `To`, following the intuition that the `To` and `Cc` mechanisms are for distinguishing roles, while all of the `Cc`'d recipients in such cases have the same role with respect to the address field.

Text processing. We tokenized and POS-tagged each email in the collection with the spaCy library [38]. In order to remove signatures, we removed all lines in an email where the majority of tokens were tagged as non-words (e.g., numbers and symbols).

A.4 Intent-tagging model

Here, we provide further details on the classifiers we used to tag the intents of emails based on email text, in our analyses of requests and commitments in Sections 4.3 and 5.2.

The models were trained on the Avocado dataset [82], and consist of recurrent neural networks with bi-directional GRU encoders using sentence-level and context-level features. The classifiers were trained on manually annotated instances with approximately 8K-10K instances per intent. The intents were annotated by human annotators who examined the entire email and determined whether or not it had a given intent. Each email instance was annotated by 3 annotators and majority voting was used to determine the final label. The Cohen's kappa coefficient for inter-annotator agreement in labeling both intents was greater than or equal to 0.61, indicating a substantial agreement among the annotators. In our experiments, we used implementations of these models that were available in our organization, Microsoft.

Received January 2020; accepted March 2020