Introduction

Many people would more readily describe a man speaking authoritatively as a *boss* than as *bossy*. With women, they'd more readily do the opposite. Women face well-documented penalties for workplace communication associated with leadership and authority. They are evaluated less favorably when they adopt dominant language (Dupree 2024), penalized for volubility that would benefit men in equivalent positions (Brescoll 2011), and sanctioned for self-promotion that is rewarded in male counterparts (Rudman 1998). In her influential work, Heilman (2001) detailed how prescriptive gender stereotypes operate to curtail women's advancement: by being subject to rigid notions of how they should behave, women who exhibit behavior typically associated with leadership are evaluated less favorably than men exhibiting the same behavior (Eagly and Karau 2002). These penalties extend to communication in particular. There is direct evidence that in the workplace women are penalized for the way they communicate (Tannen 1994; Faulkner 2009; Joshi, Wakslak, Huang, and Appel 2021).

That women face such penalties is well-established. Less clear is the channel through which gender produces them. Two accounts have been proposed. The first, grounded in role congruity theory (Eagly and Karau 2002; Rudman and Glick 2001), emphasizes gender incongruence: women are penalized when their communication violates prescriptive feminine stereotypes, such as when they adopt assertive or agentic speech. The second emphasizes status violation: women are penalized because gender shapes who is presumed entitled to claim authority, so that the same behavior is treated as legitimate from men and illegitimate from women because of men's higher ascribed status. Evidence consistent with this account includes findings that powerful women are penalized for volubility while

powerful men are not (Brescoll 2011), and that women face systematic penalties for both implicit and explicit dominance behavior (Williams and Tiedens 2016).

These channels differ along several dimensions. Gender incongruence is symmetric across audiences, since the mismatch between an actor's gender and her behavior does not depend on whom the behavior is directed toward. Status violation is directional; most acute when behavior is directed toward those of higher standing. And the channels differ in their conditionality: because status is mutable, status violation penalties may attenuate when status is legitimately earned, whereas gender incongruence carries no such conditionality, since gender does not *traditionally* change.

Distinguishing these channels empirically has proven difficult. In most organizational contexts, high-status communication is also stereotypically masculine communication: direct, assertive, abstract, or dominant. When women are penalized for such speech, the penalty is consistent with both gender incongruence and status violation. Even cases in which women strategically adapt their communication do not resolve the ambiguity: calibration toward neutral styles simultaneously avoids gender incongruence and status-claiming (Balachandra, Fischer, and Brush 2021). The two channels remain confounded whenever the behaviors under study are at once gender-typed and authority-signaling.

This study exploits a setting in which these channels come apart. In knowledge-intensive organizations, hierarchy is organized around problem solving rather than monitoring: routine problems are handled locally, while exceptional problems flow upward in the formal hierarchy (Bolton and Dewatripont 1994; Garicano 2000). When problems cross hierarchical levels, both parties must provide sufficient contextualization. Workers must convey the particulars of the problem they face, while higher-ranked actors articulate solutions and

lines of reasoning aimed at resolving it. Consequently, I argue, elaborated communication becomes locally associated with authority and leadership within such organizations, despite being stereotypically associated with feminine communication in everyday interaction (Mulac, Bradac, and Gibbons 2001). By *elaboration*, I refer to communication that makes meaning explicit through contextual detail, background information, and precise referents, rather than assuming shared understanding between speakers. A *compact* message might read "Handle this by EOD"; an elaborated version of the same directive might instead explain which client is involved, why the request is time-sensitive, what approach to prioritize, and what "this" refers to.

This association between elaboration and authority is what creates the empirical test. Because the organizational context renders elaboration a marker of status, the penalty patterns that attach to it can distinguish between the two channels, regardless of what motivates any individual speaker's use of it. If backlash is more driven by gender incongruence, women's use of a stereotypically feminine style should not be penalized, and men may face negative returns for engaging in it. If backlash is instead driven more by status violation, women should face negative returns to elaboration once it functions locally as a signal of authority, regardless of its gender-typing, while men may experience neutral or positive returns from the same behavior.

I test these predictions using two archival datasets from the same mid-sized technology firm. The first consists of multiple years of email communication between workers within the organization. In this setting, I investigate whether elaboration is indeed status-marked. I then investigate whether or not women use more elaborated communication than men, and whether women experience negative returns to elaboration in performance evaluations while

men do not. The second dataset, consisting of hiring applicant essay responses, serves as a complement to the first. Although communication in the hiring process is more formalized and ritualized than email communication, with this dataset Hypotheses 3 and 4a can be more carefully tested. Unlike in email communication, where department, rank, tasks, or other unobserved factors might render communications between men and women less comparable, the applicant essay responses are directly comparable since the communication task is identical. Further, this dataset provides an alternative setting where the outcome is different, allowing assessment of the extent to which the moderation effect of interest generalizes to broader workplace outcomes. Therefore, testing the theory in the hiring dataset constitutes an important test of Hypothesis 3, and, with regard to task uniformity, a more precise test of Hypothesis 4a.

Two additional analyses bear on that identification. If the penalty reflects status violation, it should concentrate in communication directed upward toward managers, where elaboration most clearly signals a claim to organizational authority, rather than appearing uniformly across targets. I find that it does. And if the penalty reflects status violation, it should attenuate when status is legitimately earned, since the behavior no longer constitutes an illegitimate claim. Within-person analyses support this: the penalty tracks changes in organizational standing, and among women promoted to managerial roles, it disappears. These women actually elaborate *more* to managers after promotion, yet face no consequence, because the status mismatch that generated the penalty has been resolved. Men show flat nulls throughout. Gender incongruence cannot account for this pattern, because gender does not change with promotion.

Together, these findings identify status violation as a distinct and consequential mech-

anism shaping the returns to women's workplace communication. The results reveal a paradox at the intersection of gender and organizational structure: a stereotypically feminine communication behavior, one that women adopt more readily than men, becomes penalized not because it violates gender norms but because the organizational context transforms it into an authority signal. That the penalty disappears when authority is legitimately conferred suggests it is the status mismatch, not the behavior's gender-typing, that generates the sanction. More broadly, these results imply that gender-based communication penalties are not fixed properties of particular behaviors but may track whatever styles become locally marked as signals of status. This means that organizational structures do not merely amplify existing gender disadvantages but can generate new ones by redefining which behaviors count as claims to authority.

Theory

To distinguish the two channels, we need a communication behavior that is at once gender-typed and locally associated with organizational status: the first property is what makes the channels diverge in their predictions, and the second is what lets the behavior function as a claim to authority. Linguistic elaboration is a behavior of this kind. The following subsection sets gender aside and explains why elaboration becomes salient in knowledge-intensive hierarchies and can become associated with managerial standing. The subsection that comes after that brings gender in and draws on elaboration's gender typing to separate gender incongruence from status violation.

Linguistic Elaboration in Knowledge Hierarchies

Organizations exist in large part to integrate specialized knowledge in the service of complex tasks (Grant 1996). They accomplish this integration in more than one way: through impersonal devices such as rules, sequencing, and routines, or through active exchange among actors who hold different pieces of relevant knowledge.[1] It is the latter that matters here, because it is through such exchange that communication does its integrative work. In knowledge-intensive settings especially, communication is one of the principal ways organizations make dispersed expertise usable. The knowledge-based hierarchy that Garicano (2000) develops offers a natural framework for understanding how organizations structure this exchange. In Garicano's account, workers confront problems where they arise and resolve the routine ones locally. Exceptional problems, whether technically difficult, beyond the worker's repertoire, or dependent on resources unavailable at lower levels, move upward to specialized problem solvers. In this view, hierarchy is efficient not because it allocates authority but because it economizes on problem solving and search, and channels unresolved problems to the actors whose position makes them useful sources of diagnosis, guidance, or authorization.

I adopt Garicano's framework but relax one of its stronger assumptions. Garicano takes higher-level problem solvers to possess knowledge that encompasses what the people below them already know. In many knowledge-intensive settings that assumption is too strong: frontline workers often understand the local shape of a problem better than anyone higher up, while the specialists they turn to contribute broader domain knowledge, organizational reach, or the authority to mobilize resources.[2] An exceptional problem then finds its solution not in any one person's knowledge, but in an interaction that recombines what no

single level holds on its own. This is what distinguishes *escalation* from *delegation*. In delegation, a task moves downward to someone who can already carry it out, so the knowledge the work requires already sits where it is needed. Escalation runs the opposite way and for the opposite reason: a problem moves upward precisely because the people who first encounter it cannot resolve it on their own, whether it is too technically difficult, turns on judgments about organizational interdependencies, or requires access or authorization that lower ranks lack. Because the parties to an escalation hold different and complementary pieces of what the problem requires, making themselves intelligible to one another takes work that delegation does not. Crossing ranks does not by itself raise the communicative demand; recombining complementary knowledge does.

That demand is largely linguistic. The two sides of an escalation do not share the context a problem sits in, so neither can lean on what the other already knows; each has to put into words what the other would otherwise miss. Following Bernstein (1971), I call this kind of contextualizing speech *linguistic elaboration*: communication that specifies referents, clarifies relationships, describes prior actions, and supplies enough background for someone who does not already share the immediate context to follow it. Its opposite is *linguistically compact* speech, where speakers share enough understanding to leave much unsaid. Elaboration is not length. Supplying context usually takes more words, so the two tend to move together, but what makes a message elaborated is what its language does, not how long it runs. Nor is elaboration the same as abstraction. Abstraction is a matter of how general or specific language is (Semin and Fiedler 1988; Wakslak, Smith, and Han 2014; Brysbaert, Warriner, and Kuperman 2014); elaboration is a matter of how much of the surrounding context the language makes explicit.[3]

The demand runs both ways. Upward, the message that opens an escalation has to do more than report that something has gone wrong: it has to lay out the local circumstances, name the constraints in play, and often recount what has already been tried. Downward, the reply cannot be mere shorthand either: it has to specify steps, show the reasoning behind them, and explain why a given course of action is warranted. The reply can be especially elaborated, since that is where diagnosis and the proposed solution get spelled out. The claim is comparative, not absolute: communication within a rank can be highly elaborate, but communication that crosses ranks to solve exceptional problems should call for elaboration more reliably, because the hierarchical boundary is just where shared context runs thin. Contextualizing in this way carries a quiet presumption: to supply someone with background, reasoning, and referents is to treat oneself as the party who holds what the other lacks. In an escalation that presumption is warranted on both sides, since each interlocutor genuinely commands knowledge the other needs, which is why the demand runs both ways without friction.

What makes these cross-rank exchanges demand elaboration is that they are largely diagnostic. The team-process literature separates the transition, or ideation, processes through which teams diagnose problems and search for solutions from the action processes through which they execute (Marks, Mathieu, and Zaccaro 2001; Lix, Goldberg, Srivastava, and Valentine 2022), and escalation is what pushes a problem back into that diagnostic mode: working out what is wrong and what to do about it is precisely the work that obliges each side to make its context explicit to the other.

Suggestive evidence is consistent with this view. Analyzing the Enron email corpus to find phrases that signal workplace hierarchy, Gilbert (2012) turned up a revealing pattern:

language from LIWC's "cognitive processes" category, the wording of active thinking, ran more heavily through messages sent down the hierarchy than messages sent up. I do not treat this as a direct test of elaboration. But it fits the reading that workers often reach a superior while still working a problem out rather than once they have, and that the reply is where reasoning and analysis get spelled out, which is just what one would expect if hierarchical communication carries distinctive linguistic signatures because it so often does the work of solving problems across distributed knowledge.

The managerial function at issue here is the diagnostic one, not the monitoring role that principal-agent accounts emphasize.[4] Monitoring is an important managerial function, and its importance I do not dispute. But it is a different function from the one that shapes elaboration. Managers in knowledge-intensive settings move between roles across the stages of a task: during ideation and diagnosis they act as problem solvers who clarify issues, weigh constraints, and formulate responses. Once a solution is in hand and work turns to execution, they shift toward monitoring and enforcement, and the two bleed together when an execution-phase problem itself requires escalation. The theory here concerns the problem-solving role, because that is where hierarchical communication should most shape patterns of elaboration.

The behaviors that mark status are themselves context-specific, since different production processes and authority structures elevate different behaviors to signals of standing (Magee and Galinsky 2008). Where physical labor dominates, stereotypically masculine behaviors and communication styles may stay the most salient markers (Alonso and O'Neill 2022). Knowledge-intensive organizations run on expertise, interpretive capacity, and problem-solving ability, and there elaboration becomes locally associated with status,

because cross-rank problem solving recruits it again and again and the actors higher in the hierarchy are the ones to whom exceptional problems escalate and from whom diagnostic guidance flows. The association is not merely a matter of who elaborates more often. Because contextualizing presumes that the speaker holds what the listener lacks, authoritative elaboration enacts a claim to be the more knowledgeable party, and in a setting that runs on expertise that is precisely a claim to standing. Elaboration should therefore become associated with higher organizational standing, both because higher-ranked actors do more of it and because doing it is itself a display of the standing the setting rewards.

This status association is the precondition for everything that follows: if elaboration were not status-marked in the setting, the setting would give us no ability to separate the mechanisms behind backlash against women's communication. So the first step is to establish that elaboration is associated with managerial status here, for which I use managerial rank as the observable proxy. The proxy is imperfect, and I treat the claim as probabilistic rather than exact: not every manager is the relevant problem solver in a given exchange, and non-managers often hold specialized expertise of their own. By virtue of their position, managers should on average both receive and send more elaborated messages than non-managers, since that position places them more often than others in the cross-rank exchanges where exceptional problems get contextualized and resolved. The two sides of this are not equally telling, though. That managers receive more elaborated messages shows that elaboration concentrates around managerial standing, but not, on its own, whose communication is doing the authority-marking, since the people carrying problems upward must supply context whoever they are. What managers send is the surer sign: a manager working through an escalated problem makes reasoning, constraints, and

next steps explicit, so elaboration is part of how the role itself is enacted. Sent elaboration is thus the cleaner evidence that elaboration belongs to authority, and received elaboration corroborates it.

HYPOTHESIS 1 (H1): *Controlling mechanical drivers of elaboration, managers in knowledge-intensive workplace settings will receive messages with higher elaboration than non-managers.*

HYPOTHESIS 2 (H2): *Controlling mechanical drivers of elaboration, managers in knowledge-intensive workplace settings will send messages with higher elaboration than non-managers.*

[insert Figure 1 about here]

In sum, the recurrent use of hierarchy for exceptional problem solving should make elaboration concentrate at managerial positions, not because all hierarchical communication is inherently elaborate, but because exceptional problems require actors holding different knowledge to make context explicit to one another, and managers occupy the positions through which those problems routinely pass. Delegation, monitoring, and commentary on submitted work also traverse rank, but they do not systematically demand the same contextualization. It is this coupling of elaboration with managerial position that links elaborated communication to organizational status, and that creates the setting in which the mechanisms underlying backlash against women's communication can be more cleanly distinguished.

Elaboration, Gender Typing, and Status Violation

Why does this backlash arise? Two accounts dominate, and for most of the behaviors studied they cannot be distinguished. The first, *gender incongruence*, holds that penalties follow when behavior transgresses the norms attached to a person's ascribed gender (Eagly and Karau 2002). Gender is treated as a fixed category, and the violation is a mismatch between who the actor is and how the actor behaves. The second, *status violation*, holds that penalties follow when behavior claims standing the actor has not been granted (Rudman, Moss-Racusin, Phelan, and Nauts 2012). Status, unlike gender, is positional and can be earned, conferred, or lost, so the violation is a mismatch between where the actor sits and what the behavior claims about that position. Women incur status-violation penalties in particular because gender operates as a diffuse status characteristic, one that grants men a presumption of greater worthiness than women (Ridgeway 2001).

The two accounts diverge along three dimensions, and each can be empirically exploited to determine which account is most operative in a given setting. They differ, first, in whether the actor's own gender matters. Incongruence is symmetric across gender: if a feminine-typed behavior draws a penalty because it is gender-incongruent, then men, for whom it is also incongruent, should be penalized for it too, just as men are sanctioned for communality or emotionality (Moss-Racusin, Phelan, and Rudman 2010; Rudman and Mescher 2013). Status violation works differently. Its rule is the same for everyone, in that anyone who claims standing they have not been granted invites sanction, but it does not fall evenly: because women are ascribed lower baseline status (Ridgeway 2001), the same behavior more readily reads as overreach when a woman performs it. The accounts differ, second, in whether the audience matters. A gender norm is breached the same way whoever

is watching, so incongruence predicts no penalty concentrated on any particular target.[5] A status claim exists only against a hierarchy, so status violation predicts a penalty that is sharpest when the behavior is aimed upward, at those of higher standing; overclaiming one's place draws sanction (Anderson, Ames, and Gosling 2008), while underclaiming it through humility is tolerated or even rewarded (Peters, Rowatt, and Johnson 2011). They differ, third, in whether the penalty is conditional on status. Because status can be earned, a status-violation penalty should fade once the actor legitimately holds the standing the behavior implies. Incongruence offers no such relief, because the gender category does not change with advancement, so its penalty should persist.

In most settings these accounts cannot be separated, because the behaviors usually studied are at once masculine-typed and status-marked (Tiedens 2001; Rudman et al. 2012). Anger (Brescoll and Uhlmann 2008), self-promotion (Rudman 1998), overt dominance (Williams and Tiedens 2016), dominant language (Dupree 2024): each is both, so a penalty for any of them is consistent with either account, and calibrating toward a neutral style is no more diagnostic, since it sidesteps gender incongruence and status-claiming at the same time. Telling the accounts apart requires a behavior on which they come apart, one that is stereotypically feminine yet locally marked for status. That is the behavior the first subsection identified. Because elaboration is feminine-typed, incongruence predicts no penalty for women who use it; they communicate in a gender-congruent way. Because elaboration is locally status-marked, status violation predicts the opposite, a penalty for women whose elaboration reads as a claim to authority. The two accounts now point in different directions, and if women are penalized while men are not, the penalty is hard to lay at the door of gender incongruence.[6]

Both predictions rest on a premise the argument has so far taken for granted: that elaboration is stereotypically feminine. The premise is well grounded. Across social contexts, women tend toward more elaborated speech than men, using more hedges, qualifiers, dependent clauses, and explicit referents (Hartman 1976; Crosby and Nyquist 1977; Balswick and Avertt 1977; Poole 1979; Hunt 1965; Beck 1978; Lapadat and Seesahai 1978; Mulac, Wiemann, Widenmann, and Gibson 1988; Mulac and Lundell 1994), and elaboration sits among a broader set of feminine-typed tendencies toward indirect, qualified, and affective expression, against the more direct and succinct style coded masculine.[7] These styles are not merely different; they are unequally valued. Tannen (1994) characterizes men's and women's workplace speech as nearly distinct dialects, learned through different patterns of socialization and read through different expectations, with the friction between them usually falling to women's disadvantage. Workplaces have historically privileged, and many still privilege, the style coded masculine (Faulkner 2009; Alonso and O'Neill 2022), so the gap in communicative style has been argued to feed the broader gap in women's career attainment.

A closely related and especially prominent dimension of gendered language in organizations is abstraction, and it is natural to ask whether abstraction can be used to disambiguate the mechanisms we are concerned with. Women tend to speak more concretely and less abstractly, and that tendency carries professional cost: Huang, Joshi, Wakslak, and Wu (2021) find that female entrepreneurs who use more concrete language are seen as less oriented toward growth and scalability, dampening their access to funding, and Joshi et al. (2021) argue that women's lower abstraction impedes their emergence as leaders. Yet abstraction cannot separate the two accounts, because it leaves them aligned. Abstraction

is both masculine-coded and tied to leadership, so for a woman who speaks abstractly the accounts agree rather than diverge: incongruence reads it as a woman using a masculine-coded style, status violation reads it as a status claim, and a penalty fits either. One cannot tell from such a penalty whether leadership genuinely calls for abstraction or whether abstraction reads as leaderly only because the people who have long led, men, have tended to speak that way. Elaboration is different, and that is what makes it useful. Unlike abstraction, and unlike other feminine-typed behaviors studied in organizations such as hedging, indirectness, or apologizing, which carry no offsetting authority signal, elaboration is feminine-typed while still carrying the local status association the first subsection established. For elaboration, then, the two accounts predict opposite outcomes rather than the same one.

Women who anticipate penalties for feminine-typed communication often adapt, and the adaptation carries its own costs. Women who shift toward a more masculine style under stereotype threat are rated as less warm and likeable, and evaluators grow less willing to go along with their requests (von Hippel, Wiryakusuma, Bowden, and Shochet 2011); women leaders who adopt dominant language draw backlash, more sharply for Black women (Dupree 2024); and powerful women who simply talk more than others are judged less competent and less fit to lead, while powerful men are judged the opposite (Brescoll 2011). The result is a double bind: women are penalized for communicating in stereotypically feminine ways and penalized again for adapting away from them.[8] What matters for the present argument is that this adaptation is routine in professional settings, so the gender differences in communication documented in ordinary social interaction need not carry over intact to organizations, where women are already calibrating against the expectations they

face.

Elaboration is feminine-typed, as the evidence above established, so the baseline expectation is that women elaborate more than men. The tendency is well documented across many contexts, and it is thus the natural prior for the present setting. As mentioned above, women do calibrate their behavior in professional environments, in particular, when they know that such behavior is penalized (Amanatullah and Morris 2010). This could temper the difference, but the weight of prior evidence still points toward women elaborating more. I therefore expect:

HYPOTHESIS 3 (H3): *Controlling mechanical drivers of elaboration, women in knowledge-intensive organizational settings will send messages with higher elaboration than men.*

H3 concerns who elaborates more. The question that drives the rest of the theory is a different one: not who elaborates, but how elaboration is judged when a woman uses it rather than a man.

The status account makes a concrete prediction about returns. If elaboration reads as a claim to authority when a woman uses it, then women should earn worse returns to elaboration than men, even though elaboration is, as the evidence above established, gender-congruent for women.

One qualification matters before stating this, because it governs how the prediction should be understood. To supply context is, as noted above, to presume the listener lacks what one provides, but that presumption is not always a claim to standing. When a problem is escalated, the lack of context is legitimate and the supplying of context is expected of the junior party, presuming only what the hierarchy already grants. Otherwise, elaboration may be interpreted as diagnosing, instructing, *overcontextualizing*, or otherwise claiming

the standing to speak with authority.[9] Because women are ascribed lower standing to begin with, elaboration read as a claim to authority is all the more readily policed when a woman is the one making the claim, which is what worse returns would register. With that understood, the prediction is straightforward.

HYPOTHESIS 4a (H4a): *In a knowledge-intensive workplace setting, the use of elaborated language will yield differential returns by gender, such that women will experience negative returns relative to men when using elaborated language.*

The second dimension on which the status-violation and gender-incongruence accounts diverge, the audience, refines this prediction. The talking-down reading of elaboration is not symmetric across the hierarchy. Directed downward or laterally, contextualizing presumes a knowledge gap that the speaker's standing may well warrant; directed upward, at someone whose position implies they should not need the instruction, the same elaboration most clearly presumes a standing the speaker has not been granted. A non-manager's elaboration, in other words, reads as an authority claim most saliently when it is aimed at a manager. This helps separate the authority-claim reading from a pure briefing account. A need-to-brief account can explain why upward messages require context, but it does not by itself explain why the cost should fall specifically on women or why it should attenuate once they enter the managerial category. To be clear, the prediction is not that every elaborated message sent upward draws a penalty, but that upward communication is where the ambiguity between warranted briefing and presumptive instruction is most consequential, and so where the penalty should concentrate. Incongruence, for its part, predicts no such concentration at all, since a gender norm is breached the same way whoever the message is for.

The two accounts thus make distinct, testable predictions about where the penalty falls: uniform across targets if incongruence is at work, concentrated upward if status violation is.

HYPOTHESIS 4b (H4b): *Women's negative returns to elaborated language will be concentrated in communication directed toward managers, rather than appearing uniformly across communication targets.*

The third dimension, conditionality, yields the most discriminating test. Because status can be earned, the status account predicts that the penalty should fade once a woman legitimately holds the standing her elaboration implies; incongruence predicts no such relief, since promotion does not change her gender. The cleanest way to see this is within a single person over time, following the same women as some move from non-manager to manager. As a non-manager elaborating toward those above her, a woman's elaboration reads as claiming standing she has not been granted, and the penalty should appear. Once promoted, when she elaborates toward the same class of recipients, now from within the managerial category, the claim her elaboration makes is one she is entitled to, and the penalty should attenuate. What the prediction concerns is the return to the behavior, not the amount of it. A promoted woman may well elaborate as much as before, or more; the claim is that the same elaboration, once her standing is legitimate, no longer draws the penalty it once did.

HYPOTHESIS 4c (H4c): *Within individuals, women's negative returns to manager-directed elaboration will attenuate upon promotion to managerial status, consistent with the*

conditionality of status violation penalties.

I test these predictions in two settings. The primary one is a corpus of internal email from a knowledge-intensive technology firm spanning several years, which carries all of the hypotheses. The second reexamines the gender predictions, H3 and H4a, among job applicants, where the communicative task is fixed and the same for everyone, allowing me to mitigate potential confounds on the basis of communication topic. The next section describes both settings and the measure of elaboration they share.

Empirical Setting and Measurement

These hypotheses are tested in two different contexts. First, all hypotheses are examined using a comprehensive set of email communications from a knowledge-intensive technology firm, spanning several years. This approach permits broad comparisons of language use between managers and non-managers, as well as between men and women, across an entire company and multiple time periods, in precisely the kind of setting where the main theory is expected to apply. In addition, Hypotheses 3 and 4a are evaluated again in a dataset of job applicants, where the communication under scrutiny results from a fixed task. This approach enables a reexamination of the robustness of the findings across varying contexts and outcomes, and allows for a direct comparison of returns to language use when the task is identical for both genders.

Data Description

The first empirical setting consists of the complete email corpus of a mid-sized technology company, spanning the years 2007-2017, as well as the company's human resource personnel records. The company was structured into several departments, of which three are explicitly identified in the data: tech, marketing, and sales. The email data set consists of over 11 million distinct messages, including meta-data (anonymized sender and receiver identification codes, timestamps, etc.) and content. The personnel data includes employee age, tenure, gender, manager versus non-manager status, and attainment outcomes such as performance evaluations and departure information. Additional preprocessing was conducted on the data for the purpose of this analysis. For instance, the data were segmented into monthly and yearly intervals, structural measures such as network centrality were computed for any given employee at any month in the observation window, and emails crossing department boundaries or hierarchical ranks were identified. See Table 1 for summary statistics of the email corpus.

[insert Table 1 about here]

The second data sample, pertaining to hiring, is composed of job candidate application information from a mid-sized technology firm that was hiring between 2015 and 2016. The firm requested writing samples to essay questions from applicants during the evaluation process. The sample incorporates the optional essay responses, application records including applicant information, and human resource records. In total, these data contain sociodemographic information of the applicants, including gender[10]; anonymized identification codes; information regarding the positions they applied to; information regarding the quality of the organizations on their resumes; information on whether or not they were

hired; and information regarding who they were interviewed by (among other details). The full sample consisted of 13,254 job applicants, of which 2,128 were interviewed by 161 distinct interviewers, and only 358 were hired. The gender split was 8,739 men and 4,290 women.

The essay responses in the hiring dataset are answers to three questions that were designed to assess candidate fit with the company along the dimensions of ideal workplace, lifestyle preferences, and ideological orientation. The responses were collected over the course of 18 months. The essay responses were optional, and Stein, Goldberg, and Srivastava (2018) find that they are the result of 17% of 79,192 total job applicants deciding to provide a response. This does introduce selection concerns, but according to their analysis, in terms of observable characteristics, applicants who chose to provide responses and those who did not are remarkably similar. For example, the proportion of women who chose to answer versus not-answer were near identical (answerers were 32% female and non-answerers were 33% female); average ages were similar (25 for answerers and 27 for non-answerers); and measures of human capital were similar. The one major difference they find is that respondents who chose not to answer were much more likely to have been sourced through a recruiter. Because the analyses here turn on comparisons by gender, and the responding and non-responding pools were nearly identical in gender composition, this source-of-recruitment difference is unlikely to bias the comparisons of interest.

[insert Table 2 about here]

The two datasets are analyzed separately, in separate archival "studies", yet they reinforce each other in important ways. The first dataset contains email communications between and across hierarchical rank, and between and across genders, over the course of

several years. It thus provides rich and direct visibility into cross-rank communications, which allow proper assessment of the first two hypotheses. Additionally, it offers a crucial advantage over previous studies of gender and workplace communication: it captures relatively naturalistic communication patterns. While workplace email maintains some degree of formality, it represents more routine, day-to-day interactions compared to the highly ritualized contexts examined in prior work, such as investor pitches or formal presentations. This setting reveals how gender shapes communication when speakers have more discretion in their linguistic choices, rather than conforming to rigid contextual scripts. As a result, both genuine gender differences in communication style and their consequences for workplace outcomes can be better identified.

The second dataset, consisting of hiring applicant essay responses, serves as a complement to the first. Although communication in the hiring process is more formalized and ritualized than email communication, with this dataset Hypotheses 3 and 4a can be more carefully tested. Unlike in email communication, where department, rank, tasks, or other unobserved factors might render communications between men and women less comparable, the applicant essay responses are directly comparable since the communication task is identical. Further, this dataset provides an alternative setting where the outcome is different, allowing assessment of the extent to which the moderation effect of interest generalizes to broader workplace outcomes. Therefore, testing the theory in the hiring dataset constitutes an important test of Hypothesis 3, and, with regard to task uniformity, a more precise test of Hypothesis 4a. A further reason makes this setting well suited to Hypothesis 3 in particular. Calibration operates where people know which styles a setting penalizes, and applicants, writing their way into a firm whose local norms they have no

way to know, have had no such occasion to calibrate. The baseline gender difference in elaboration should therefore be more detectable among them with less of the dampening that an embedded workforce may have already applied.

Processing the Email Corpus

The email corpus is organized into folders of communication associated with each de-identified sender who was a member of the organization and sent an email during the observation period. A given sender folder may have hundreds or even thousands of email files. Each email file contains content data and meta-data. The meta-data pertain to recipient information (including who the message may have been forwarded or copied to), subject information, and date and time information. These are intersected with human resources data to produce a more complete view of the network of communication. From the human resources data sociodemographic information such as age and gender are matched to senders and receivers, as well as information pertaining to the rank of the interlocutors (i.e., whether they are a manager or not) and their departments (marketing, tech, sales, or neither).

To the email content, two models are applied, in order to extract two distinct measures. The first is *linguistic elaboration*, a measure of how implicit or explicit the communication in the email is, computed with the Stacked Ensemble Model of Elaborated Language (SEMEL). This measure and its associated model are described further in Appendix C. A second model, BERTopic (Grootendorst 2022), was applied in order to extract a measure of the primary topic discussed in the email. The BERTopic model was trained on a random 45% sample of the full email corpus (roughly 5 million emails) in order to categorize

the emails by topic. BERTopic is a topic modeling technique that leverages BERT embeddings and class-based-TF-IDF to create dense clusters allowing for easily interpretable topics whilst keeping important words in the topic descriptions. The advantage of the transformers-based embedding is that semantic meaning of the topic is encoded, and so text can be matched to topics by virtue of distance in the semantic space, even when no words in the text directly match the topic's key terms. The topics generated from BERTopic were used to identify emails which pertain to non-work content and exclude them from the elaboration computations. This allows comparison of discourse that pertains only to work-related phenomena between men, women, managers, and non-managers (although the proportion of personal topics discussed are included as controls in the main models)[11].

The measure of linguistic elaboration is aggregated to the month and year level for two reasons. First, the human resources data are organized by month intervals, and so the two sets can be joined much more naturally when both are organized that way. Second, an outcome of interest for Hypothesis 4a – *performance rating* – is annual. Thus, when estimating determinants of that outcome, it is regressed against data aggregated to the annual level. Linguistic elaboration is aggregated for a specific sender by simply taking the average of each score across all that sender's emails, within a given month. To get to the annual measure, those elaboration scores are instead averaged over the given year.

The *performance rating* of employees is used to assess Hypothesis 4a and is set to 1 for employees whose underlying score indicated effective performance and set to 0 otherwise. It is binarized because the company's rating system changed during the observation period. For more details, see Goldberg, Srivastava, Manian, Monroe, and Potts (2016).

A slew of control variables are constructed in order to account for confounding, reduce omitted variable bias, and mitigate potential endogeneity. For each control variable, information is either sourced from the human resources data or the variable is computed using the patterns of email communication. When a given control variable is calculated in this manner, it is constructed at both the month and year level, so that analyses can be conducted at both levels.

The first control variable is the *network centrality* of the sender, and is based on the individual's eigenvector centrality. It is included in order to account for structural factors which may impact the amount of context the employee is privy to, as well as evaluation outcomes, since central actors are known to be influential (Bonacich 1987; Rossman, Esparza, and Bonacich 2010). Next, the analysis follows Goldberg et al. (2016) and further accounts for structural position of each individual within the communication network of the company, by assessing the redundancy of their local network. This is quantified using Burt's (1992) measure of *network constraint*. The network constraint measure reveals the extent to which a worker is limited in their access to novel information, which is relevant to the degree of contextualization they would employ in discourse. Burt's constraint is defined as:

$$B_i = \sum_j \left(w_{ij} + \sum_k w_{ik} w_{kj} \right)^2, i \neq k \neq j \tag{1}$$

Where B_i is Burt's constraint on worker i, w_{ij} is the proportion of worker i's communication activity spent directly on worker j, and the inner summation captures the proportion

of worker i's network activity spent indirectly on worker j. In addition, the *tenure* of the worker (length of time in the organization), their *age*, and their *department* (sales, tech, marketing, or neither) are all accounted for by accessing human resources records.

Next, potential "mechanical drivers" of elaboration are accounted for, such as the duration of the email thread or its depth. These mechanical drivers are features of the conversation, such as how long it's been going on, that determine the amount of explicitness necessary for effective communication. The first of these is *average reception depth*, which refers to the average depth in an email chain in which a worker is first tagged, either via direct addressing or by being copied (these are accounted for separately, as they are considered indicative of separate types of communication). "Depth" refers to the average number of messages exchanged pertaining to the same subject prior to the focal worker being tagged. Average reception depth is a driver of linguistic elaboration because context is created and established during an exchange, so the level of mutual compactness will vary over the duration of the conversation. This is no different in email threads, which may first start out establishing the details of an issue and by the end progress to much more implicit communication styles. Similarly, the *duration* of the thread is measured in units of time. This is to account for the probability that context about the subject is exchanged outside of email, which increases the longer the thread is alive, even if the number of messages exchanged within it are low.

The next mechanical drivers concern the volume of emails and exchange partners a given worker handles. These are *count of emails received*, *count of emails sent*, and *unique exchange partners*. Finally, linguistic elaboration is examined strictly for those emails that concern work-related topics, the proportion of a worker's communication that concern

personal topics is also accounted for. This is done because it may indicate familiarity, which may also affect linguistic compactness. Personal communication is captured with two variables, *proportion of personal topics received* and *proportion of personal topics sent*.

Processing the Hiring Data Corpus

The hiring corpus is organized as a file of applicant information, human resources information, and essay responses linked by applicant identification code. Similarly to the email corpus, both SEMEL and BERTopic are applied to the essay responses, in order to extract measures of linguistic elaboration as well as response topics. For the response topics, rather than categorize them by "personal" versus "work-related", the full set of topics identified by the model is retained, and these are used as controls in the analysis, in order to account for differences in actual content associated with each applicant's response.

For this setting the determinants of two outcomes are estimated. The first is *mean elaboration*, which is the average elaboration of the respondent across all three essay response questions. In the analysis for which this outcome is based, the difference in elaboration between men and women is measured, conditional on all observable drivers of elaboration. The second dependent variable of interest is *rejected*, which is an indicator of whether or not the applicant was hired. It is coded such that it takes a value of 1 if the applicant is not hired, and 0 otherwise. In this analysis, differences in gender-based returns to elaboration are explored in terms of probability of rejection.

Following Stein et al. (2018), a host of control variables related to the applicant characteristics and application details are accounted for. The first is *managerial role*, which is an indicator of whether or not the applicant applied for a managerial role. Secondly, *age* is

controlled for, which was imputed based on the assumption that the oldest date listed on a candidate's resume is likely to be their date of college graduation. This was assumed to correspond to an age of 22 at the time. As Stein et al. (2018) note, this approach introduces measurement error, but in any case serves as a useful enough proxy, and furthermore captures overall work experience which is of core relevance to hiring. The next two variables are *top university* and *top organization*, and are used as controls since these are known to be attractive signals of quality for job candidates. *Top university* is determined by whether or not the applicant listed any top 20 educational institution on their resume, as defined by the 2017 U.S. News rankings. *Top organization* is determined similarly, and set to 1 if a candidate listed any of Fortune's 2017 "Most Admired Companies" on their resumes. Additionally, *job source* is controlled for, which is the channel through which the candidate discovered the job posting. These include recruiter, referral, company website, job board, university job fair, or "other". The final variable is *log word count*. This control captures to some extent levels of motivation and effort, which are relevant to hiring decisions.

Measuring Linguistic Elaboration

The crux of the analysis turns on the measure of linguistic elaboration. Elaboration, in this context, is derived from a definition first articulated by Bernstein (1971), which characterizes utterances in terms of the extent to which the language used implies taken-for-granted information between speakers. In the measure derived from this framework, language that is descriptive, precise, and explicit is considered *elaborated*. Elaborated code is employed when speakers cannot assume full background contextual knowledge on the part of the listener and therefore must articulate meaning carefully and provide context. By contrast,

language that eschews description, is imprecise, or remains implicit is considered *compact*. Compact code is employed when there is little need to supply context or description because meaning is assumed to be apparent. Of course, there is no sharp distinction between elaborated and compact language, and the same utterance may be elaborate in some respects while compact in others. Even so, the concept is useful as a general characterization of expressions, and is informative about the extent to which speakers assume shared context on a given topic.

[insert Table 3 about here]

Table 3 makes the construct more concrete. All examples are initial emails in a thread, which helps hold constant differences that might otherwise arise from message position within an ongoing exchange and thus makes the table more informative about communicative tendencies rather than conversational location. Across the higher-elaboration examples, communication tends to supply more of the relevant organizational context rather than presuming it: actors are identified, prior actions are described, constraints are made explicit, and rationale or next steps are spelled out for the recipient. The lower-elaboration examples, by contrast, more often state the immediate request, update, or technical issue more directly while relying more heavily on shared situational understanding. These differences are sometimes subtle to the unaided eye, which is precisely why a dedicated measure is useful. But they are not arbitrary: across cases, the score recurrently tracks the degree to which communication makes context explicit for an audience that cannot be assumed to share it fully. The examples in Table 3 are chosen to make this contrast legible, while Appendix D provides a fuller set of examples showing that both managers

and non-managers produce communication at higher and lower levels of elaboration across technical, sales, and internal operational contexts.

Linguistic elaboration is measured by applying SEMEL to all utterances in the corpora. SEMEL is an ensemble model that combines multiple natural language processing approaches to extract features from a text and encode them into a machine-interpretable format. The constituents of the ensemble include part-of-speech tagging, syntactic parsing (Jurafsky and Martin 2020), word embedding (Bojanowski, Grave, Joulin, and Mikolov 2016), co-reference resolution (Jurafsky and Martin 2020), formality assessment (Heylighen and Dewaele 1999), readability assessment (Coleman and Liau 1975), type-to-token computation, assessment of term rarity, and word counting. These features are treated as relevant to appraising elaboration and are encoded as representations in a feature matrix.

Once encoded, the final stage of the ensemble is quantification, whereby a dense neural network receives the matrix of features and outputs an elaboration score based on the encoded features. A more detailed description of the encoding and quantification process can be found in Appendix C.

SEMEL is designed to capture the kind of contextual explicitness illustrated in Table 3. Syntactic and coreference features help distinguish language that makes referents, relationships, and prior actions explicit from language that leaves more of that context implicit, while lexical diversity, formality, and related features capture additional dimensions of elaboration. In external validation on the Wizard of Wikipedia dialogues (Dinan, Roller, Shuster, Fan, Auli, and Weston 2019), where one speaker is tasked with supplying knowledge and the other with seeking it, SEMEL assigns higher elaboration to the knowledge-providing role, consistent with the construct's emphasis on context-supplying

communication (see Appendix C for more details). In the email corpus, elaboration is only moderately correlated with word count ($r = 0.27$), indicating that the measure captures linguistic qualities beyond message length.

Estimation Procedure

In the email corpus, several sets of models are estimated corresponding to the different hypotheses and dependent variables. To test Hypotheses 1, 2, and 3, where the outcome is *linguistic elaboration*, fixed-effects regressions are estimated with month-fixed effects. To test Hypothesis 4a, where the outcome is *performance rating*, linear probability models with year-fixed effects are estimated. The interaction between gender and elaboration is the coefficient of interest. In all between-person email models, standard errors are clustered by individual to account for within-person serial correlation across observation periods.

Hypothesis 4b decomposes the elaboration penalty by communication target. For non-managers of each gender, separate regressions estimate the returns to elaboration directed toward male colleagues, female colleagues, managers, and cross-department recipients. Because these target categories reflect overlapping organizational relationships, each is estimated in a separate model.

Hypothesis 4c employs individual fixed-effects models that absorb all time-invariant characteristics of the sender, ruling out the possibility that the penalty reflects selection on stable unobservables. Two specifications are estimated. The first pools all person-years for each gender, regressing performance rating on elaboration directed toward managers, with own manager status included as a control. This specification captures the broad trajectory of status evolution within the organization, including informal changes in standing

that do not correspond to formal promotion. The second restricts the sample to individuals observed in both non-manager and manager person-years and includes an interaction between own manager status and manager-directed elaboration. This transitioner specification directly tests whether the returns to upward-directed elaboration change upon promotion, trading statistical power for greater precision in isolating the discrete moment at which formal authority is conferred. In both specifications, standard errors are clustered by individual.

Using the hiring corpus, Hypotheses 3 and 4a are evaluated. The first set of models, testing Hypothesis 3, are fixed-effects regressions with interviewer-fixed effects, topic-fixed effects, and heteroscedasticity consistent standard errors, where the outcome is *linguistic elaboration*. The second set, testing Hypothesis 4a, are generalized linear models with a logistic link function, since the outcome is *rejected*, a binary indicator of whether or not the applicant was rejected. These models include interviewer-fixed effects and topic-fixed effects, with heteroscedasticity consistent standard errors.

Results

The results are organized by hypothesis. Each hypothesis is evaluated using evidence from the email corpus, the hiring corpus, or both, depending on the data available for the relevant test. Tables 1 and 2 present descriptive statistics for the email and hiring datasets, respectively.

H1-H2: Elaboration and Hierarchical Status

Table 4 displays the results of three model estimations showing the relationship between hierarchical rank and mean elaboration of messages received. The output variable and the predictors are standardized. The models in this table test Hypothesis 1, the prediction that managers in knowledge-intensive workplace settings will receive messages with higher elaboration than non-managers. All three models support the hypothesis. Model 1 estimates the relationship accounting for network position, age, and tenure. In this model, the largest effect is manager status, which is associated with a 0.191 standard deviation increase in linguistic elaboration of messages received. Manager status remains positively associated with elaboration received when department is accounted for (Model 2), and when mechanical drivers of elaboration are accounted for (Model 3).

[insert Table 4 about here]

[insert Figure 2 about here]

Similarly, Table 5 displays the results of three model estimations showing the relationship between hierarchical rank and mean elaboration of messages sent, corresponding to Hypothesis 2. All three models support the hypothesis. Across both tables, manager status has a larger effect size than the network measures and a more consistent positive association. Figure 2 displays the distributions of elaboration in emails sent from managers compared to emails sent from non-managers. Managers exhibit moderately but systematically higher elaboration in their communication. The peak of the managers' distribution (0.498) is separated from the non-managers' peak (0.465) by approximately one-third of a standard deviation, with the distributions crossing at elaboration levels of 0.459. Beyond this point,

elaborated communication becomes increasingly characteristic of managerial status. Figure 3 shows the estimated relationship between rank and elaboration of messages sent.

[insert Table 5 about here]

Together, these results establish that linguistic elaboration is associated with managerial status in this setting. The communication managers produce is the more direct evidence that elaboration is authority-marked, since elaborating is part of how the role is enacted, while the communication they receive corroborates that elaboration concentrates around managerial standing. This association is the precondition for the mechanism-identification tests that follow: if elaboration were not status-marked, the setting would offer no leverage for distinguishing gender incongruence from status violation as explanations for backlash against women's communication.

[insert Figure 3 about here]

H3: Gender and Elaboration

Table 6 estimates the relationship between gender and linguistic elaboration of sent messages in the email corpus, corresponding to Hypothesis 3. Model 7 controls for network position, tenure, and age; Model 8 adds department and manager status; Model 9 adds mechanical drivers of elaboration. Across all three specifications, the coefficient on gender is positive, indicating that women use somewhat more elaborated language than men. However, the effect is modest in magnitude and does not reach conventional levels of statistical

significance in any specification ($\beta = 0.070$, $p = 0.25$ in the baseline; $\beta = 0.047$, $p = 0.50$ in the full model).

This pattern is consistent with the theoretical expectation that while women tend toward more elaborated speech in general, strategic calibration in organizational settings may attenuate the baseline difference. As discussed in the theory section, women in professional contexts are known to adjust their communication in response to workplace norms, and the modest email corpus results may reflect such adjustment.

[insert Table 6 about here]

The same prediction is tested among job applicants to the firm, whose essay responses were written during the hiring process (see Table 2 for descriptive statistics). These essay responses normalize the task and topic of communication, which helps control for factors that might independently drive the level of elaboration men and women produce. Unlike the email setting, where employees have had time to learn which communication styles are locally rewarded or penalized, applicants in the hiring process have not yet been embedded in the organization. If the baseline gender difference in elaboration is real but attenuated by communication calibration, it should be more clearly visible in this setting where the motivation to calibrate toward local norms is more limited. Table 7 presents these results. In both specifications, women use significantly more elaborated language than men ($\beta = 0.025$, $p = 0.01$ in Model 10; $\beta = 0.023$, $p < 0.05$ in Model 11 with topic fixed effects), even after controlling for essay length, applicant characteristics, and application source.

[insert Table 7 about here]

Taken together, the two datasets paint a consistent picture. Women do tend toward more elaborated communication, as the sociolinguistic literature predicts. The difference is clearly present among job applicants who have not yet adapted to the organizational environment, and attenuated among employees who have. This pattern provides qualified support for H3 while also illustrating the calibration dynamic that previous research has demonstrated to be at play.

H4a: Gender-Differential Returns to Elaboration

Figure 4 displays results from the analysis of Hypothesis 4a, which predicts that the use of elaborated language will yield differential returns by gender, such that women will experience negative returns relative to men. The visual is stark: for women, the probability of a favorable rating begins dropping around elaboration levels of 0.45, almost exactly where managerial and non-managerial communication patterns diverge (Figure 2). For men, the probability of a favorable rating mildly increases with increased elaboration. Table 8 presents the corresponding model estimates. The interaction between gender and elaboration is negative and significant ($\beta = -0.076$, $p = 0.009$), indicating that women experience substantially worse returns to elaboration than men. The main effect of elaboration is positive ($\beta = 0.021$, $p = 0.043$), meaning that for men, increased elaboration is associated with modestly better performance evaluations. For women, the combined effect ($0.021 - 0.076 = -0.055$) implies that elaboration is penalized.

[insert Figure 4 about here]

[insert Table 8 about here]

The hiring corpus provides a replication of this test in a setting where the communication task is fixed, meaning no task differences between men and women can account for any differences in observed returns to elaboration. Table 9 presents the interaction between gender and elaboration on the probability of an applicant being rejected. The interaction is positive and significant in Model 13 ($\beta = 0.120$, $p = 0.036$), with the topic-fixed-effects marginally significant in Model 14 ($\beta = 0.120$, $p = 0.057$). These results indicate that elaboration increases rejection probability for women relative to men. Figure 5 plots this interaction. As the value of elaboration increases, the predicted probability of rejection decreases substantially for men while it slightly increases for women. That this pattern replicates across two distinct professional contexts (albeit within the same firm), and two different outcomes provides converging support for H4a.

[insert Figure 5 about here]

[insert Table 9 about here]

To further assess the mechanisms, the email corpus interaction is decomposed by the focal employee's hierarchical status. Recall that H4a is motivated by the status violation account, which predicts that penalties should concentrate among non-managers, for whom elaboration constitutes an illegitimate claim to authority. Gender incongruence, by contrast, predicts penalties regardless of rank, since gender does not change with position. Table 10 presents these results. Among non-managers, the penalty is large and significant: women face a -0.105 penalty relative to men in Model 15 ($p < 0.05$) and -0.103 in Model 16 ($p < 0.01$), while the positive main effect of elaboration ($\beta = 0.034$, $p < 0.01$ in Model

15) indicates that men non-managers are rewarded for the same behavior. Among managers, the interaction is effectively zero ($\beta = 0.003$ and 0.006, both nonsignificant), and the main effect of elaboration is likewise null. Figure 6 visually depicts these findings.

[insert Figure 6 about here]

[insert Table 10 about here]

These results are consistent with the status violation account. Across both genders, a managerial role confers sufficient status to render elaboration unremarkable; no penalty or reward attaches to it. Among non-managers, however, the status congruence between speech and speaker depends on gender. Men non-managers appear to be rewarded for displays of authority-marked communication, while women non-managers are penalized for it. This pattern is difficult to reconcile with a simple gender-incongruence account. Linguistic elaboration is stereotypically feminine, yet women are penalized for using it while men experience neutral or positive returns. Instead, the findings suggest that gender constrains who is afforded latitude to signal authority through communication style.

Although not reaching conventional levels of statistical significance, post-hoc analyses of the hiring data reveal a pattern that aligns with this finding. For non-managerial positions, elaboration appears marginally beneficial overall ($\beta = -0.109$, $p < 0.10$), but women face incrementally higher rejection probabilities as elaboration increases (interaction $\beta = 0.100$, $p = 0.11$). For managerial positions, neither the main effect of elaboration ($\beta = -0.095$, $p = 0.52$) nor the interaction ($\beta = 0.218$, $p = 0.17$) reaches significance. That the pattern of results in the hiring data mirrors the email corpus findings, despite narrowly falling short of conventional significance thresholds, provides suggestive cross-context evidence for the

role of hierarchical status in shaping gender-based returns to elaboration.

H4b: Concentration of Penalties in Upward Communication

Hypothesis 4b predicts that women's negative returns to elaboration will concentrate in communication directed toward managers, rather than appearing uniformly across communication targets. If the penalty reflects gender incongruence, it should appear regardless of audience, since the mismatch between the actor's gender and her behavior does not depend on whom the behavior is directed toward. If it reflects status violation, it should be most acute in upward communication, where elaboration most clearly signals a claim to organizational authority.

To test this prediction, elaboration effects are decomposed by recipient type. Specifically, four distinct communication patterns are analyzed for non-managers of each gender: elaboration in messages directed toward male colleagues, female colleagues, managers, and cross-department recipients. Because these categories reflect overlapping organizational relationships, separate estimations are conducted for each target type. Table 11 presents these results, Figure 7 displays them.

[insert Figure 7 about here]

Among women non-managers, the penalty is large and significant only when elaboration is directed toward managers ($\beta = -0.088$, $p < 0.01$).[12] Elaboration directed toward male colleagues ($\beta = -0.010$, $p = 0.89$), female colleagues ($\beta = -0.052$, $p = 0.42$), and cross-department recipients ($\beta = -0.025$, $p = 0.68$) yields no significant penalty. Among men

non-managers, men face no penalty for elaboration in any direction. They are positively rewarded for elaborating to male colleagues ($\beta = 0.027$, $p < 0.001$) and cross-department ($\beta = 0.016$, $p < 0.05$), while elaboration toward female colleagues ($\beta = 0.008$, $p = 0.65$) and managers ($\beta = 0.003$, $p = 0.77$) is null. These results support H4b.

The comparison across genders is informative beyond the hypothesis itself. Men's lateral elaboration is rewarded while women's is not, suggesting that displaying knowledge through elaboration benefits men even in contexts where no status violation is at play. Women who elaborate laterally are neither rewarded nor penalized; the behavior is tolerated but not credited. Directing elaboration upward eliminates the reward for men (from positive to null) and introduces a penalty for women (from null to negative). Both genders thus appear to bear some cost for upward-directed elaboration, but from different baselines: men lose a reward, women incur a penalty.

However, these between-person comparisons leave open the possibility that the localization reflects selection rather than a causal effect of communication context: women who direct more elaborated communication to managers may differ from other women in unobserved ways that independently affect their evaluations. The next analysis addresses this concern.

[insert Table 11 about here]

H4c: Status Conditionality

Hypothesis 4c draws on the conditionality of status violation. Because status is mutable, the penalty for upward-directed elaboration should attenuate as a woman's organizational

status increases. Gender incongruence predicts no such attenuation, since the gender category does not change with status advancement. A within-person design allows this prediction to be tested by holding constant all stable characteristics of the individual and asking whether the returns to manager-directed elaboration vary as the sender's position in the organization evolves over time.

Table 12 presents individual fixed-effects models for women and men separately, pooling across all person-years and using elaboration directed toward managers as the key predictor, controlling for own manager status. By absorbing all stable individual characteristics, these models rule out the possibility that the penalty reflects selection: it is not the case that a particular type of woman who elaborates toward managers is independently evaluated less favorably. At the same time, these models capture the effect of time-varying changes in a woman's organizational standing, including but not limited to formal promotions. As a woman becomes more established, more widely recognized as authoritative, or more central to the organization's problem-solving processes, the status mismatch between her position and her use of authority-marked communication narrows. If the penalty reflects that mismatch, it should weaken as her status rises. For women, the within-person coefficient is negative and significant ($\beta = -0.133$, $p < 0.05$), confirming that the penalty operates within individuals over time and is not an artifact of selection. For men, the coefficient is near zero ($\beta = -0.027$, $p = 0.54$). Notably, the magnitude of the women's penalty is attenuated relative to the within-person estimate of -0.173 obtained when restricting the sample to non-manager person-years, consistent with the penalty concentrating in periods of lower organizational status. The estimation sample is reduced to 149 women and 261 men person-years due to the intersection of non-missing elaboration-to-manager scores,

performance ratings, and the full set of covariates.[13] These results support H4c.

[insert Table 12 about here]

A follow-up analysis isolates one specific and observable form of status change: formal promotion from non-manager to manager. The sample is restricted to 58 women and 108 men who are observed in both non-manager and manager person-years during the panel. This restriction reduces power in exchange for greater precision of the estimated effect. For each gender, an individual fixed-effects model is estimated with elaboration directed toward managers, own manager status, and their interaction. The interaction directly tests whether the returns to manager-directed elaboration change upon promotion. Table 13 presents these results.

Among women, the base effect of manager-directed elaboration when serving as a non-manager is large and negative ($\beta = -0.249$, $p < 0.01$). The interaction with own manager status is positive and nearly equal in magnitude ($\beta = 0.278$, $p = 0.065$), yielding a net effect when serving as a manager of essentially zero ($-0.249 + 0.278 = 0.029$). The penalty for upward-directed elaboration disappears entirely upon promotion. Among men, there is no penalty in either condition: the base effect is null ($\beta = 0.061$, $p = 0.41$), the interaction is null ($\beta = -0.036$, $p = 0.60$), and the net effect as manager is likewise null (0.025).

[insert Table 13 about here]

The pattern is exactly what status violation predicts. When the same woman lacks formal authority and elaborates upward, she is penalized. When she gains legitimate sta-

tus through promotion and communicates the same way to the same class of recipients, the penalty vanishes. Men never face a penalty in either condition. The interaction term for women is marginally significant at conventional levels, reflecting the limited degrees of freedom available from 51 person-years after missingness. However, the direction, magnitude, and near-perfect offset of the base and interaction coefficients constitute a pattern that is difficult to attribute to noise. Notably, the descriptive data show that women transitioners actually elaborate *more* to managers after promotion (mean elaboration of 0.163 as managers versus 0.059 as non-managers), yet the penalty disappears. They engage in more of the penalized behavior and face no consequence, because the status mismatch that generated the penalty has been resolved.

Taken together, the pooled within-person analysis and the transitioner follow-up provide convergent evidence that the penalty for women's upward-directed elaboration is conditional on the sender's organizational status. The pooled analysis captures the broad trajectory of status evolution within the organization; the transitioner analysis isolates the discrete moment at which formal authority is conferred. Both point to the same conclusion: when the status mismatch is present, the penalty operates, and when it is resolved, the penalty disappears. Gender incongruence cannot account for this pattern, because gender does not change with status advancement.

Discussion and Conclusions

The Channel Through Which Gender Operates

This study set out to distinguish two channels through which gender might produce backlash against women's workplace communication: gender incongruence, in which penalties arise because behavior violates prescriptive gender norms, and status violation, in which penalties arise because behavior claims authority that the actor is not presumed to hold. By exploiting a setting in which a stereotypically feminine communication behavior, linguistic elaboration, becomes locally associated with authority through its role in cross-rank problem solving, I am able to separate these channels empirically. The pattern that emerges is consistent. Elaboration is status-marked in knowledge-intensive organizations: managers send and receive more elaborated communication than non-managers. Women use more elaborated communication than men, yet experience negative returns to elaboration while men do not. That women are penalized for gender-congruent communication, and that men, for whom elaboration is gender-incongruent, face no penalty at all, is difficult to square with a simple gender-incongruence account. The penalty does not track the violation of gender norms; it tracks who is presumed entitled to claim authority, which is gender operating through *status ascription* rather than *behavioral prescription*. Two further results point to status violation as that channel: the penalty concentrates in upward communication directed toward managers, and it disappears among women promoted to managerial roles, even as those women elaborate more after promotion. The same behavior that marks a woman as *bossy* when she lacks formal authority marks her as a *boss* once that authority is conferred.

Conditionality and the Within-Person Evidence

The within-person analyses provide a distinct form of evidence for the status violation interpretation, by ruling out the selection concerns that between-person comparisons cannot. Between-person comparisons, however carefully controlled, leave open the possibility that women who elaborate toward managers differ in unobserved ways from those who do not. The individual fixed-effects models eliminate this concern by comparing the same woman to herself across periods in which her organizational standing changes. The penalty persists within individuals, confirming it is not a selection artifact. And it attenuates as organizational standing rises: the pooled within-person estimate is smaller in magnitude than the non-manager-only estimate, and among women who transition to managerial roles, the penalty disappears.

This has direct implications for organizational practice. The concentration in upward communication indicates that the penalty is produced by a specific set of actors: those occupying positions of authority who are evaluating communication from below. It is managers receiving upward-directed elaboration from women who register it as an illegitimate status claim. This suggests that interventions targeting how authority figures evaluate communication from subordinates are better aligned with the source of the penalty than efforts aimed at changing how women communicate or at shifting gender norms across the organization as a whole.

Organizational Structures as Sources of New Disadvantage

Prior work on gender and organizations has largely focused on how organizational structures amplify existing gender disadvantages. The present results suggest that organizations

also generate new ones. In knowledge-intensive settings, the production process creates a status marker, linguistic elaboration, that would not carry authority connotations elsewhere. Because elaboration is stereotypically feminine, women use it more readily than men, and because the organizational context renders it a signal of authority, their natural communication tendencies become a liability. The disadvantage is not imported; it is produced endogenously by the intersection of organizational structure and gender. The lateral asymmetry reinforces this: men are rewarded for the same behavior that is merely tolerated in women, suggesting that the status channel operates through differential reward as well as through penalty.

This implies that gender-based communication penalties are not fixed properties of particular behaviors but track whatever styles become locally marked as signals of status. Organizations that restructure their knowledge flows or alter the communicative demands of cross-rank interaction may inadvertently create new status markers and, with them, new gendered penalties.

Limitations, Scope Conditions, and Generalizability

Both datasets are from the same knowledge-intensive technology firm, where problem solving and information sharing are central to work processes. The relationship between elaboration and status identified here depends on that organizational context. In settings where hierarchy rests more strongly on physical output, credentials, or seniority rather than on knowledge contribution, different behaviors become status-marked, and elaboration in particular need carry no status charge at all; the specific penalty documented here may not apply. This points to a distinction between two claims of different scope. The specific

finding, that linguistic elaboration carries a gendered penalty, is bounded to settings where elaboration is the locally status-marked behavior. The general claim is mechanism-level and travels further: wherever a behavior becomes a local marker of authority, those lacking ascribed standing should face penalties for displaying it, and where that behavior is also gender-typed, the penalty will fall unevenly by gender. What this study contributes is identification of that mechanism in a case where it can be cleanly observed, not a claim that elaboration specifically is penalized everywhere.

The study relies on archival communication data and cannot observe the evaluative process through which elaboration is interpreted. The identification strategy does not require such observation. The theoretical predictions concern outcome patterns, not perceptions, and the evidence is structured to distinguish the two channels through those patterns. Two specific alternative explanations were also ruled out. The penalty does not reflect women being more prone to problems requiring elaborate explanation: women receive slightly less negative feedback than men, and men experience stronger evaluation penalties for negative feedback (see Figure A1 and Table A1 in Appendix A). Nor does it reflect gendered sanctions for inefficient communication: elaboration and word count are only moderately correlated ($r = 0.27$), word count is positively associated with performance ratings, and the gender interaction on word count is not significant (see Figure B1 and Table B1 in Appendix B).

The within-person analyses and the hiring data address the two most serious threats to causal identification. The individual fixed-effects models rule out selection on stable unobservables by showing that the penalty operates within individuals over time. The hiring data rule out feedback loops between past evaluations and communication styles by

demonstrating the same gender-based penalty in initial candidate evaluations, where no prior organizational history exists. Together, these designs substantially narrow the set of confounds that could account for the observed pattern.

Notes

[1]These devices are not mutually exclusive. Impersonal rules and directives standardize interaction (Van De Ven, Delbecq, and Koenig 1976), sequencing structures the flow of information across interdependent tasks, and routines sustain coordination without continuous explicit instruction (Winter 1986). Even where coordination is partly routinized, these mechanisms still depend on communication and common knowledge (Nonaka and Takeuchi 1995; Grant 1996), including the transactive processes through which workers learn who knows what (Wegner 1987; Huber and Lewis 2010).

[2]This adaptation is consistent with Grant (1996)'s conception of the knowledge-based firm, in which specialized knowledge can reside in different individuals independently of their formal position in the hierarchy. Higher-level actors here know enough to be useful problem solvers, but they need not fully subsume the local knowledge that frontline workers hold.

[3]To illustrate this difference, consider an example from Wakslak et al. (2014). As they note, in describing an earthquake, one might state that 120 people died and 400 were injured (a concrete statement conveying specific details) or that the earthquake is a national tragedy (an abstract statement conveying higher-level meaning). On the elaborateness/compactness continuum, one might also state that 120 people died and 400 were injured (relatively compact) or that 120 *Americans tragically* died, and 400 were *gravely* injured (relatively more elaborate). These versions differ in elaboration, but both are concrete. Similarly, one could say that the earthquake is a national tragedy (relatively compact) or that the earthquake is an *American* tragedy (relatively more elaborate). Here, the versions likewise differ in elaboration, but this time they are both abstract. The illustration uses added words for clarity, but elaboration is not merely a matter of adjectives, emotional coloring, or descriptive detail. More fully, it is the supplying of context, including referents, relationships, and prior actions, that an audience cannot be assumed to share.

[4]Principal-agent accounts cast managers as monitors who curb agency problems and secure adequate effort (Alchian and Demsetz 1972; Jensen and Meckling 1976); see also Lazear (1986, 2000).

[5]In fact, Moss-Racusin et al. (2010) find that men who present modestly are evaluated more harshly than comparably modest women, a policing that tends to be done by other men (Kimmel 1994). On an incongruence account, then, men have at least as much reason as women to be penalized for a feminine-typed behavior.

[6]The identification here turns on women being penalized for a behavior that is gender-congruent for them while men are not comparably penalized for the same behavior; had men faced systematic penalties for elaboration, the reading would differ.

[7]Researchers comparing communication across gender repeatedly find that men and women draw on different genres of speech and different skills for "doing things with words" (Harding 1975). The differences fall along recognizable lines: feminine communication tends to be more indirect, elaborate, and affective, where "affective" concerns the internal states of persons and "instrumental" the external attributes of objects, while masculine communication tends to be more direct, succinct, and instrumental (Mulac et al. 2001; Popp, Donovan, Crawford, L, Marsh, and Peele 2003; Newman, Groom, Handelman, and Pennebaker 2008; von Hippel et al. 2011).

[8]Calibration can sometimes succeed. In entrepreneurial pitching, Balachandra et al. (2021) show that women pare back most feminine-typed features of their speech while keeping emotional expressiveness, a careful recalibration rather than wholesale adoption of a masculine style, and Zorn and Violanti (1996) find that once women adopt standardized styles, their communication ability predicts career attainment about as well as men's. Even so, adaptation does not reliably shield women from penalty.

[9]See Hall (1976, p. 92)'s account of how extensive contextualization may be interpreted as presumptive or as a form of "talking-down" (perhaps especially when directed upward), given that contextualizing itself presumes the speaker has knowledge their fellow interlocutor does not.

[10]For a proportion of the applicants (13%) it was necessary to apply an imputation process to assess their gender, using names and year of birth. See Stein et al. (2018) for details.

[11]On average, topics identified as "personal" constitute only about 2% of emails sent and received by employees in the firm (see Table 1 for further reference).

[12]This result survives Benjamini-Hochberg correction across the four women's target-type tests ($p_{BH} = 0.036$).

[13]The primary source of missingness is the performance rating outcome, which is available for only a subset of person-years. Elaboration to managers is additionally missing for person-years in which the individual did not email a manager.

References

Alchian, Armen A. and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization." *The American Economic Review* 62:777–795.

Alonso, Natalya M. and Olivia (Mandy) O'Neill. 2022. "Going Along to Get Ahead: The Asymmetric Effects of Sexist Joviality on Status Conferral." *Organization Science* 33:1794–1815.

Amanatullah, Emily T. and Michael W. Morris. 2010. "Negotiating Gender Roles: Gender Differences in Assertive Negotiating Are Mediated by Women's Fear of Backlash and Attenuated When Negotiating on Behalf of Others." *Journal of Personality and Social Psychology* 98:256–267.

Anderson, Cameron, Daniel R. Ames, and Samuel D. Gosling. 2008. "Punishing Hubris: The Perils of Overestimating One's Status in a Group." *Personality and Social Psychology Bulletin* 34:90–101.

Balachandra, Lakshmi, Katrin Fischer, and Candida Brush. 2021. "Do (Women's) Words Matter? The Influence of Gendered Language in Entrepreneurial Pitching." *Journal of Business Venturing Insights* 15:e00224.

Balswick, J and C P Avertt. 1977. "Differences in Expressiveness: Gender, Interpersonal Orientation, and Perceived Parental Expressiveness as Contributing Factors." *Journal of Marriage and the Family* 39:121–127.

Beck, R. 1978. "Sex Differentiated Speech Codes." *International Journal of Women's Studies* 1:566–572.

Bernstein, Basil. 1971. *Class, Codes and Control: Volume 1 - Theoretical Studies Towards A Sociology Of Language*. Routledge Kegan Paul Ltd.

Bojanowski, Piotr, Edouard Grave, Armand Joulin, and Tomás Mikolov. 2016. "Enriching Word Vectors with Subword Information." *CoRR* abs/1607.04606.

Bolton, Patrick and Mathias Dewatripont. 1994. "The Firm as a Communication Network." *The Quarterly Journal of Economics* 109:809–839.

Bonacich, Phillip. 1987. "Power and Centrality: A Family of Measures." *American Journal of Sociology* 92:1170–1182.

Brescoll, Victoria L. 2011. "Who Takes the Floor and Why: Gender, Power, and Volubility in Organizations." *Administrative Science Quarterly* 56:622–641.

Brescoll, Victoria L. and Eric Luis Uhlmann. 2008. "Can an Angry Woman Get Ahead? Status Conferral, Gender, and Expression of Emotion in the Workplace." *Psychological Science* 19:268–275.

Brysbaert, M., A. B. Warriner, and V. Kuperman. 2014. "Concreteness Ratings for 40 Thousand Generally Known English Word Lemmas." *Behavior Research Methods* 46:904–911.

Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Coleman, Meri and T. L. Liau. 1975. "A Computer Readability Formula Designed for Machine Scoring." *Journal of Applied Psychology* 60:283–284.

Crosby, F. and L. Nyquist. 1977. "The Female Register: An Empirical Study of Lakoff's Hypothesis." *Language in Society* 6:313–322.

Davies, Mark. 2008. "The Corpus of Contemporary American English (COCA): 560 Million Words, 1990-Present." *www.english-corpora.org/coca/* .

Dinan, Emily, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. 2019. "Wizard of Wikipedia: Knowledge-Powered Conversational Agents." *arXiv preprint arXiv:1811.01241* .

Dupree, Cydney H. 2024. "Words of a Leader: The Importance of Intersectionality for Understanding Women Leaders' Use of Dominant Language and How Others Receive It." *Administrative Science Quarterly* 69:271–323.

Eagly, Alice H. and Steven J. Karau. 2002. "Role Congruity Theory of Prejudice Toward Female Leaders." *Psychological Review* 109:573–598.

Faulkner, Wendy. 2009. "Doing Gender in Engineering Workplace Cultures: Observations from the field." *Engineering Studies* 1:3–18.

Francis, W. Nelson and Henry Kucera. 1979. *Brown Corpus Manual: Manual of Information to Accompany a Standard Corpus of Present-Day Edited American English for Use with Digital Computers*. Providence, RI: Brown University Department of Linguistics.

Garicano, Luis. 2000. "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy* 108:874–904.

Gilbert, Eric. 2012. "Phrases That Signal Workplace Hierarchy." In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, pp. 1037–1046, New York, NY, USA. Association for Computing Machinery.

Goldberg, Amir, Sameer B. Srivastava, V. Govind Manian, William Monroe, and Christopher Potts. 2016. "Fitting In or Standing Out? The Tradeoffs of Structural and Cultural Embeddedness." *American Sociological Review* 81:1190–1222.

Grant, Robert M. 1996. "Toward a Knowledge-Based Theory of the Firm." *Strategic Management Journal* 17:109–122.

Grootendorst, Maarten. 2022. "BERTopic: Neural Topic Modeling with a Class-based TF-IDF Procedure." *arXiv preprint arXiv:2203.05794* .

Hall, Edward T. 1976. *Beyond Culture*. Anchor Press.

Harding, Susan. 1975. "Women and Words in a Spanish Village." In *Toward an Anthropology of Women*, edited by Rayna R. Reiter, pp. 283–308. Monthly Review Press.

Hartman, M. 1976. "A Descriptive Study of the Language of Men and Women born in Maine around 1900 as it Reflects the Lakoff Hypotheses in Language and Woman's Place." In *The Sociology of the Languages of American Women*, edited by Betty Lou Dubois and Isabel Crouch, pp. 81–90. San Antonio, TX: Trinity University Press.

Heilman, Madeline E. 2001. "Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder." *Journal of Social Issues* 57:657–674.

Heylighen, Francis and Jean-Marc Dewaele. 1999. "Formality of Language: Definition, Measurement and Behavioral Determinants." *Internal Report, Center "Leo Apostel", Free University of Brussels* http://pespmc1.vub.ac.be/Papers/Formality.pdf.

Huang, Laura, Priyanka Joshi, Cheryl Wakslak, and Andy Wu. 2021. "Sizing Up Entrepreneurial Potential: Gender Differences in Communication and Investor Perceptions of Long-Term Growth and Scalability." *Academy of Management Journal* 64:716–740.

Huber, George P. and Kyle Lewis. 2010. "Cross-Understanding: Implications for Group Cognition and Performance." *Academy of Management Review* 35:6–26.

Hunt, K. W. 1965. *Grammatical Structures Written at Three Grade Levels*. Champaign, IL: National Council of Teachers of English.

Hutto, C.J. and Eric Gilbert. 2014. "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text." In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, pp. 216–225.

Jensen, Michael C. and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3:305–360.

Joshi, Priyanka D., Cheryl J. Wakslak, Laura Huang, and Gil Appel. 2021. "Gender Differences in Communicative Abstraction and their Organizational Implications." *Rutgers Business Review* 6:145–153.

Jurafsky, Dan and James H. Martin. 2020. *Speech and Language Processing*. (3rd ed. draft).

Kimmel, Michael. 1994. "Masculinity as Homophobia: Fear, Shame, and Silence in the Construction of Gender Identity." In *Theorizing Masculinities*, edited by Harry Brod and Michael Kaufman, pp. 119–141. SAGE Publications.

Lapadat, J. and M. Seesahai. 1978. "Male versus Female Codes in Informal Contexts." *Sociolinguistics Newsletter* 8:7–8.

Lazear, Edward P. 1986. "Salaries and Piece Rates." *Journal of Business* 59:405–431.

Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review* 90:1346–1361.

Lewis, David D, Yiming Yang, Tony Rose, and Fan Li. 2004. "RCV1: A New Benchmark Collection for Text Categorization Research." *Journal of Machine Learning Research* 5:361–397.

Lix, Katharina, Amir Goldberg, Sameer B Srivastava, and Melissa A Valentine. 2022. "Aligning Differences: Discursive Diversity and Team Performance." *Management Science* 68:8430–8448.

Magee, Joe C. and Adam D. Galinsky. 2008. "Social Hierarchy: The Self-Reinforcing Nature of Power and Status." *Academy of Management Annals* 2:351–398.

Marks, Michelle A., John E. Mathieu, and Stephen J. Zaccaro. 2001. "A Temporally Based Framework and Taxonomy of Team Processes." *The Academy of Management Review* 26:356–376.

Moss-Racusin, Corinne A., Julie E. Phelan, and Laurie A. Rudman. 2010. "When Men Break the Gender Rules: Status Incongruity and Backlash Against Modest Men." *Psychology of Men & Masculinity* 11:140–151.

Mulac, Anthony, James J. Bradac, and Pamela Gibbons. 2001. "Empirical Support for the Gender-as-Culture Hypothesis: An Intercultural Analysis of Male/Female Language Differences." *Human Communication Research* 27:121–152.

Mulac, Anthony and Torborg Louisa Lundell. 1994. "Effects of Gender-Linked Language Differences in Adults' Written Discourse: Multivariate Tests of Language Effects." *Language Communication* 14:299–309. Special Issue Language Attitutes.

Mulac, Anthony, John M. Wiemann, Sally J. Widenmann, and Toni W. Gibson. 1988. "Male/Female Language Differences and Effects in Same-Sex and Mixed-Sex Dyads: The Gender-Linked Language Effect." *Communication Monographs* 55:315–335.

Newman, Matthew L., Carla J. Groom, Lori D. Handelman, and James W. Pennebaker. 2008. "Gender Differences in Language Use: An Analysis of 14,000 Text Samples." *Discourse Processes* 45:211–236.

Nonaka, Ikujiro and Hirotaka Takeuchi. 1995. *The Knowledge Creating Company*. New York: Oxford University Press.

Peters, Amber S., Wade C. Rowatt, and Megan K. Johnson. 2011. "Associations Between Dispositional Humility and Social Relationship Quality." *Psychology* 2:155–161.

Piantadosi, Steven T. 2014. "Zipf's Word Frequency Law in Natural Language: A Critical Review and Future Directions." *Psychonomic Bulletin & Review* 21:1112–1130.

Poole, Millicent E. 1979. "Social Class, Sex, and Linguistic Coding." *Language and Speech* 22:49–67.

Popp, Danielle, Roxanne A. Donovan, Mary Crawford, Katherine L, Kerry L. Marsh, and Melanie Peele. 2003. "Gender, Race, and Speech Style Stereotypes." *Sex Roles* 48:317–325.

Ridgeway, Cecilia L. 2001. "Gender, Status, and Leadership." *Journal of Social Issues* 57:637–655.

Rossman, Gabriel, Nicole Esparza, and Phillip Bonacich. 2010. "I'd Like to Thank the Academy, Team Spillovers, and Network Centrality." *American Sociological Review* 75:31–51.

Rudman, Laurie A. 1998. "Self-Promotion as a Risk Factor for Women: The Costs and Benefits of Counterstereotypical Impression Management." *Journal of Personality and Social Psychology* 74:629–645.

Rudman, Laurie A and Peter Glick. 2001. "Prescriptive Gender Stereotypes and Backlash toward Agentic Women." *Journal of Social Issues* 57:743–762.

Rudman, Laurie A. and Kris Mescher. 2013. "Penalizing Men Who Request a Family Leave: Is Flexibility Stigma a Femininity Stigma?" *Journal of Social Issues* 69:322–340.

Rudman, Laurie A., Corinne A. Moss-Racusin, Julie E. Phelan, and Sanne Nauts. 2012. "Status Incongruity and Backlash Effects: Defending the Gender Hierarchy Motivates Prejudice Against Female Leaders." *Journal of Experimental Social Psychology* 48:165–179.

Semin, G. R. and K. Fiedler. 1988. "The Cognitive Functions of Linguistic Categories in Describing Persons: Social Cognition and Language." *Journal of Personality and Social Psychology* 54:558–568.

Sichel, H. S. 1975. "On a Distribution Law for Word Frequencies." *Journal of the American Statistical Association* 70:542–547.

Stein, Sarah Kathryn, Amir Goldberg, and Sameer B Srivastava. 2018. "Cultural Matching in the Hiring Process: Evidence from Online Job Postings." *Stanford University, Graduate School of Business Research Papers* .

Tannen, Deborah. 1994. *Talking from 9 to 5: How Women's and Men's Conversational Styles Affect Who Gets Heard, Who Gets Credit, and What Gets Done at Work*. New York: W. Morrow.

Tiedens, Larissa Z. 2001. "Anger and Advancement Versus Sadness and Subjugation: The Effect of Negative Emotion Expressions on Social Status Conferral." *Journal of Personality and Social Psychology* 80:86–94.

Van De Ven, Andrew H., Andre L. Delbecq, and Richard Koenig. 1976. "Determinants of Coordination Modes Within Organizations." *American Sociological Review* 41:322–338.

von Hippel, Courtney, Cindy Wiryakusuma, Jessica Bowden, and Megan Shochet. 2011. "Stereotype Threat and Female Communication Styles." *Personality and Social Psychology Bulletin* 37:1312–1324.

Wakslak, C. J., P. K. Smith, and A. Han. 2014. "Using Abstract Language Signals Power." *Journal of Personality and Social Psychology* 107:41–55.

Warstadt, Alex, Amanpreet Singh, and Samuel R Bowman. 2018. "Neural Network Acceptability Judgments." *arXiv preprint arXiv:1805.12471* .

Wegner, Daniel M. 1987. "Transactive Memory: A Contemporary Analysis of the Group Mind." In *Theories of Group Behavior*, edited by Brian Mullen and George R Goethals, volume 9, pp. 185–208. Springer New York.

Williams, Melissa J. and Larissa Z. Tiedens. 2016. "The Subtle Suspension of Backlash: A Meta-Analysis of Penalties for Women's Implicit and Explicit Dominance Behavior." *Psychological Bulletin* 142:165–197.

Winter, Sidney G. 1986. "The Research Program of the Behavioral Theory of the Firm: Orthodox Critique and Evolutionary Perspective." In *Handbook of Behavioral Economics*, edited by Benjamin Gilad and Stanley Kaish, volume A, pp. 151–188. Greenwich, CT: JAI Press.

Wolf, Thomas. 2017. "State-of-the-Art Neural Coreference Resolution for Chatbots." *HuggingFace Blog: https://medium.com/huggingface/state-of-the-art-neural-coreference-resolution-for-chatbots-3302365dcf30* .

Zorn, Theodore E and Michelle T Violanti. 1996. "Communication Abilities and Individual Achievement in Organizations." *Management Communication Quarterly* 10:139–167.

Tables

Table 1: Summary Statistics of Variables (Email Corpus)

Variable	N	Mean	Std. Dev.	Min	Median	Max
Mean Month Elab. Sending	24,397	0.472	0.112	0.000	0.469	0.979
Network Constraint (std.)	43,232	0.000	1.000	-0.713	-0.324	7.664
Centrality (std.)	39,207	0.000	1.000	-0.823	-0.285	16.566
Tenure (months)	44,138	15.95	16.32	1.00	10.00	110.00
Age	31,302	34.06	10.02	17.12	31.26	76.50
Mean Convo Rec Depth (monthly)	16,330	176.43	1,025.14	1.00	42.64	34,562.00
Count Emails Received (monthly)	16,339	616	1,022.76	1	385	58,001
Count Emails Sent (monthly)	24,397	412.5	2,073.76	1.0	131.0	258,333.0
Mean Convo CC Depth (monthly)	31,216	88.67	904.67	1.00	12.65	58,302.00
Unique Exchange Initiators (monthly)	16,333	59.07	47.71	1.00	52.00	356.00
Proportion PT Emails Sent	10,275	0.02	0.051	0.00	0.01	1.00
Month Proportion PT Emails Rec	11,197	0.02	0.033	0.00	0.01	1.00
Subject Duration Tagged (days)	16,330	128.7	124.0	0.0	99.9	1,314.1
Manager (Dummy)	3,981	-	-	-	-	-
Sales (Dummy)	11,263	-	-	-	-	-
Marketing (Dummy)	5,707	-	-	-	-	-
Woman (Dummy)	9,936	-	-	-	-	-

Table 2: Summary Statistics of Hiring Data

Variable	N	Mean	Std. Dev.	Min	Median	Max
Mean Elaboration Score	13,254	0.5603	0.1461	0.0518	0.5611	0.9788
Mean Word Count	13,254	51.77	38.99	2.00	42.67	513.00
Variable	Counts (13,254)					
Woman	4,290			-		
Top Organization Flag	1,549			-		
Top University Flag	5,784			-		
Manager Role Applied	3,154			-		
Variable (Job Source)	Counts (13,254)					
Career Site	6,383			-		
Employee Referral	521			-		
Internal	85			-		
Recruiter	47			-		
University	33			-		
Job Board	6,123			-		
Other	62			-		

Table 3: Illustrative Examples of Higher and Lower Linguistic Elaboration

Example	Sender status	SEMEL	Excerpt	Key feature
Panel A. Higher-elaboration examples				
H1	Manager	.901	"[NAME1] reached out to discuss ways she could contribute to the services program . . . at present, the team does not have a full-time member representing the voice of the customer . . . I suggested that [NAME1] shadow the team over the next couple of weeks, and we can regroup in early February to discuss next steps."	Actors, background, current gap, proposed next step.
H2	Manager	.901	"I spoke with [NAME1], and we will only need to supply public IPs for [LOCATION1]; the remaining locations are out of scope . . . internal private DHCP addresses are not relevant for this requirement . . . please provide the list of public static IP addresses that should be authorized."	Scope clarified; policy rationale and request explicit.
H3	Manager	.889	"The phones are compatible with multiple backend systems . . . the problem appears to be local to the temporary space . . . humidity, room temperature swings, and voltage drops could all be contributing factors . . . we are now awaiting new power supplies and battery systems in an effort to troubleshoot the issue further."	Problem framed; possible causes and troubleshooting described.
Panel B. Lower-elaboration examples				
L1	Non-manager	.440	"Would you mind listening to the entire call before giving me feedback suggesting that I may not have done so . . . I feel like I pushed that point as far as I could, but not every customer responds to it . . . perhaps we can work on ways for me to do that even more effectively."	Immediate exchange emphasized; less broader contextual scaffolding.
L2	Non-manager	.413	"This project had some complicated inspection call-outs by the AHJ . . . it took some time for the team to work through the design-change options and coordinate with the customer . . . the work is scheduled to be completed today, with inspection either today or tomorrow."	Issue stated more directly, with less explicit organizational framing.
L3	Non-manager	.447	"From the photos, it does not look like the wire is stranded . . . if it is [TECH1], it is out of compliance, but if it is [TECH2], we are fine . . . based on the information available, this should be acceptable, and I will revise the plan set to reflect the new information."	Narrow technical issue resolved with less broader situational context.

Notes: Excerpts are lightly edited for readability and anonymized. Placeholder indices reset within each email excerpt. All examples are initial emails in a thread. Examples are selected for illustrative contrast and therefore pair higher-elaboration examples with managers and lower-elaboration examples with non-managers, although both managers and non-managers produce communication spanning the full range of elaboration in the data. A fuller set of examples appears in Appendix D.

Table 4: Relationship between Hierarchical Rank and Linguistic Elaboration of Messages Received (Standardized)

Variable	Model 1	Model 2	Model 3
Manager	0.191***	0.162***	0.154***
	(0.000)	(0.000)	(0.000)
Network Constraint (std.)	0.078***	-0.028	-0.093
	(0.004)	(0.244)	(0.101)
Network Centrality (std.)	-0.034*	0.016	0.093***
	(0.081)	(0.214)	(0.000)
Tenure (std.)	0.046***	0.009	0.015
	(0.008)	(0.523)	(0.342)
Age (std.)	0.094***	0.067***	0.084***
	(0.000)	(0.000)	(0.000)
Sales		-0.475***	-0.345***
		(0.000)	(0.000)
Marketing		-0.122**	-0.030
		(0.012)	(0.668)
Tech		-0.029	-0.102*
		(0.575)	(0.089)
Avg Rec Thread Depth (std.)		0.120***	0.171***
		(0.000)	(0.000)
Avg CC Thread Depth (std.)		-0.069*	-0.096
		(0.079)	(0.125)
Count Emails Sent (std.)			-0.010*
			(0.076)
Count Emails Received (std.)			0.196***
			(0.000)
Unique Exchange Initiators			-0.355***
			(0.000)
Prop. Personal Topics Sent (std.)			-0.038***
			(0.004)
Prop. Personal Topics Rec. (std.)			-0.038
			(0.209)
Mean Subj. Duration (std.)			-0.223***
			(0.000)
Period Fixed-Effects	Yes	Yes	Yes
N	13,262	12,397	5,297
R^2	0.0266	0.1289	0.3095
Adj. R^2	0.0262	0.1282	0.3074

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Standard errors clustered by individual

Table 5: Relationship between Hierarchical Rank and Linguistic Elaboration of Messages Sent (Standardized)

Variable	Model 4	Model 5	Model 6
Manager	0.249***	0.226***	0.253***
	(0.000)	(0.002)	(0.000)
Network Constraint (std.)	0.047	0.033	0.106*
	(0.112)	(0.279)	(0.099)
Network Centrality (std.)	0.077***	0.102***	0.182***
	(0.000)	(0.000)	(0.000)
Tenure (std.)	0.015	-0.003	0.015
	(0.612)	(0.924)	(0.588)
Age (std.)	0.050*	0.013	0.007
	(0.065)	(0.611)	(0.819)
Sales		-0.341***	-0.329***
		(0.000)	(0.000)
Marketing		-0.006	0.040
		(0.937)	(0.692)
Tech		0.074	0.116
		(0.330)	(0.205)
Avg Rec Thread Depth (std.)		-0.016*	-0.004
		(0.093)	(0.809)
Avg CC Thread Depth (std.)		-0.082	-0.100
		(0.133)	(0.178)
Count Emails Sent (std.)			0.047**
			(0.013)
Count Emails Received (std.)			-0.007
			(0.837)
Unique Exchange Initiators			-0.074**
			(0.049)
Prop. Personal Topics Sent (std.)			-0.060**
			(0.012)
Prop. Personal Topics Rec. (std.)			-0.058***
			(0.003)
Mean Subj. Duration (std.)			0.007
			(0.860)
Period Fixed-Effects	Yes	Yes	Yes
N	11,022	10,615	5,298
R^2	0.0253	0.0598	0.1170
Adj. R^2	0.0249	0.0589	0.1143

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Standard errors clustered by individual

Table 6: Relationship between Gender and Linguistic Elaboration of Messages Sent (Standardized)

Variable	Model 7	Model 8	Model 9
Woman	0.070	0.046	0.047
	(0.250)	(0.441)	(0.502)
Network Constraint (std.)	0.045*	0.035	0.106*
	(0.073)	(0.256)	(0.099)
Network Centrality (std.)	0.095***	0.102***	0.185***
	(0.000)	(0.000)	(0.000)
Tenure (std.)	0.041	-0.001	0.017
	(0.146)	(0.962)	(0.588)
Age (std.)	0.061**	0.013	0.005
	(0.027)	(0.623)	(0.819)
Manager		0.227***	0.257***
		(0.001)	(0.000)
Sales		-0.336***	-0.321***
		(0.000)	(0.000)
Marketing		-0.011	0.036
		(0.887)	(0.692)
Tech		0.077	0.120
		(0.311)	(0.205)
Avg Rec Thread Depth (std.)		-0.016*	-0.004
		(0.094)	(0.809)
Avg CC Thread Depth (std.)		-0.082	-0.099
		(0.135)	(0.178)
Count Emails Sent (std.)			0.047**
			(0.013)
Count Emails Received (std.)			-0.007
			(0.837)
Unique Exchange Initiators			-0.076**
			(0.049)
Prop. Personal Topics Sent (std.)			-0.061**
			(0.012)
Prop. Personal Topics Rec. (std.)			-0.058***
			(0.003)
Mean Subj. Duration (std.)			0.005
			(0.860)
Period Fixed-Effects	Yes	Yes	Yes
N	12,095	10,615	5,298
R^2	0.0192	0.0604	0.1177
Adj. R^2	0.0188	0.0594	0.1149

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Standard errors clustered by individual

Table 7: Relationship between Gender and Linguistic Elaboration of Essay Responses (Standardized)

Variable	**Model 10**	**Model 11**
Woman	0.025***	0.023**
	(0.010)	(0.018)
Seeking Management Role	0.008	0.006
	(0.446)	(0.549)
Age (std.)	-0.012***	-0.012***
	(0.010)	(0.007)
Top Organization	0.061***	0.062***
	(0.000)	(0.000)
Top University	-0.005	-0.005
	(0.559)	(0.577)
Recruited	0.038	0.017
	(0.718)	(0.871)
Referred	0.172***	0.157**
	(0.006)	
Company Website	0.132**	0.119**
	(0.024)	(0.041)
Job Board	0.093	0.080
	(0.109)	(0.170)
Other Job Source	0.179**	0.171**
	(0.042)	(0.048)
University Job Fair	0.285***	0.273***
	(0.006)	(0.005)
Log Word Count (std.)	0.882***	0.876***
	(0.000)	(0.000)
Topic Fixed-Effects	No	Yes
N	10,389	10,389
R^2	0.7754	0.7760
Adj. R^2	0.7751	0.7756

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Robust standard errors in parentheses

Table 8: Interaction between Gender and Linguistic Elaboration of Messages Sent on Probability of Favorable Rating (Standardized)

Variable	Model 12
Woman × Elaboration Used (std.)	-0.076***
	(0.009)
Woman	-0.008
	(0.634)
Elaboration Used (std.)	0.021**
	(0.043)
Network Constraint (std.)	0.036
	(0.472)
Network Centrality (std.)	-0.068
	(0.368)
Tenure (std.)	0.035***
	(0.000)
Age (std.)	-0.043**
	(0.029)
Manager	0.058***
	(0.003)
Sales	-0.070
	(0.315)
Marketing	0.016
	(0.750)
Tech	-0.044
	(0.365)
Avg Rec Thread Depth (std.)	0.003
	(0.516)
Avg CC Thread Depth (std.)	0.026
	(0.375)
Count Emails Sent (std.)	0.004
	(0.788)
Count Emails Received (std.)	0.023**
	(0.016)
Unique Exchange Initiators	0.029
	(0.434)
Prop. Personal Topics Sent (std.)	0.002
	(0.930)
Prop. Personal Topics Rec. (std.)	0.001
	(0.973)
Period Fixed-Effects	Yes
N	6,209
R^2	0.0505
Adj. R^2	0.0470

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Standard errors clustered by individual

Table 9: Interaction between Gender and Linguistic Elaboration of Essay Responses on Probability of Rejection (Standardized)

Variable	Model 13	Model 14
Woman × Elaboration Used (std.)	0.120**	0.120**
	(0.036)	(0.057)
Woman	0.322	-0.139
	(0.549)	(0.545)
Elaboration Used (std.)	-0.075	-0.526
	(0.193)	(0.184)
Seeking Management Role	0.792***	0.791***
	(0.000)	(0.000)
Age (std.)	-0.056**	-0.011**
	(0.013)	(0.014)
Top Organization	-0.095	-0.096
	(0.232)	(0.232)
Top University	-0.034	-0.036
	(0.519)	(0.500)
Recruited	1.137*	1.143*
	(0.081)	(0.079)
Referred	0.212	0.221
	(0.591)	(0.576)
Company Website	0.338	0.345
	(0.362)	(0.353)
Job Board	0.724*	0.730*
	(0.052)	(0.050)
Other Job Source	0.665	0.677
	(0.250)	(0.242)
University Job Fair	1.294*	1.304*
	(0.077)	(0.075)
Log Word Count (std.)	0.215***	0.282***
	(0.000)	(0.000)
Interviewer Fixed-Effects	Yes	Yes
Topic Fixed-Effects	No	Yes
N	10,389	10,389
R^2	0.0833	0.0835
Adj. R^2	0.0710	0.0706

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Robust standard errors in parentheses

Table 10: Interaction between Gender and Linguistic Elaboration of Messages Sent on Probability of Favorable Rating, by Hierarchical Status (Standardized)

Variable	Non-Managers Model 15	Non-Managers Model 16	Managers Model 17	Managers Model 18
Woman × Elaboration Used (std.)	-0.105**	-0.103***	0.003	0.006
	(0.0417)	(0.0372)	(0.0143)	(0.0177)
Woman	-0.019	-0.012	-0.014	-0.026
	(0.0121)	(0.0239)	(0.0230)	(0.0297)
Elaboration Used (std.)	0.034***	0.024**	-0.013	-0.011
	(0.0091)	(0.0119)	(0.0126)	(0.0100)
Network Constraint (std.)	0.026	0.018	-0.024	-0.017
	(0.0200)	(0.0583)	(0.0859)	(0.1002)
Network Centrality (std.)	-0.011	-0.114	0.028	0.023**
	(0.0317)	(0.0838)	(0.0232)	(0.0099)
Tenure (std.)	0.079***	0.067***	-0.010	-0.013
	(0.0155)	(0.0107)	(0.0124)	(0.0148)
Age (std.)	-0.051**	-0.050**	-0.040	-0.035
	(0.0201)	(0.0228)	(0.0318)	(0.0274)
Sales	-0.068	-0.056	-0.109	-0.118*
	(0.0877)	(0.0869)	(0.0726)	(0.0712)
Marketing	0.023	0.037	-0.029	-0.039
	(0.0423)	(0.0573)	(0.0439)	(0.0429)
Tech	-0.043	-0.090	0.004	-0.051*
	(0.0303)	(0.0578)	(0.0259)	(0.0262)
Avg Rec Thread Depth (std.)	0.002	0.000	-0.004	-0.001
	(0.0057)	(0.0057)	(0.0038)	(0.0067)
Avg CC Thread Depth (std.)	0.008	0.023	0.088	0.099
	(0.0238)	(0.0269)	(0.1432)	(0.1298)
Count Emails Sent (std.)		-0.008		-0.002
		(0.0130)		(0.0032)
Count Emails Received (std.)		0.033		0.017
		(0.0247)		(0.0138)
Unique Exchange Initiators		0.049		-0.011
		(0.0471)		(0.0082)
Prop. Personal Topics Sent (std.)		0.017		-0.003
		(0.0149)		(0.0401)
Prop. Personal Topics Rec. (std.)		0.006		0.094**
		(0.0299)		(0.0478)
Period Fixed-Effects	Yes	Yes	Yes	Yes
N	5,593	5,191	1,058	1,018
R^2	0.0499	0.0616	0.0847	0.1064
Adj. R^2	0.0470	0.0576	0.0697	0.0866

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Standard errors clustered by individual

Table 11: Returns to Elaboration by Communication Target and Gender, Non-Managers (Standardized)

Target Type	**Women Non-Managers**	**Men Non-Managers**
Elaboration to Men	-0.010	0.027***
	(0.8905)	(0.0000)
Elaboration to Women	-0.052	0.008
	(0.4204)	(0.6467)
Elaboration to Managers	-0.088***	0.003
	(0.0089)	(0.7716)
Elaboration Cross-Department	-0.025	0.016**
	(0.6760)	(0.0346)
Controls	Yes	Yes
Period Fixed-Effects	Yes	Yes
N (To Men)	1,609	3,582
N (To Women)	1,609	3,576
N (To Managers)	1,017	1,930
N (Cross-Department)	1,609	3,582

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Standard errors clustered by individual. Each row is a separate regression of favorable rating on elaboration directed toward the indicated target type, with full controls for network position, tenure, age, department, email volume, and mechanical drivers of elaboration.

Table 12: Within-Person Returns to Manager-Directed Elaboration on Probability of Favorable Rating, All Person-Years (Standardized)

Variable	Women Model 19	Men Model 20
Elaboration to Managers (std.)	-0.133**	-0.027
	(0.0187)	(0.5445)
Manager	-0.031	-0.053
	(0.7465)	(0.5809)
Network Constraint (std.)	0.209	0.179*
	(0.4083)	(0.0756)
Network Centrality (std.)	-0.018	-0.099
	(0.8615)	(0.2705)
Tenure (std.)	0.226***	0.197***
	(0.0031)	(0.0009)
Avg Rec Thread Depth (std.)	-0.016	0.010
	(0.7073)	(0.5526)
Avg CC Thread Depth (std.)	-0.168	0.140
	(0.1260)	(0.3053)
Count Emails Received (std.)	0.030	-0.028
	(0.3725)	(0.2677)
Count Emails Sent (std.)	0.002	0.021
	(0.7379)	(0.3673)
Unique Exchange Initiators	0.004	0.108***
	(0.9076)	(0.0037)
Prop. Personal Topics Sent (std.)	0.650**	0.046
	(0.0169)	(0.7336)
Prop. Personal Topics Rec. (std.)	0.082	0.059
	(0.5517)	(0.5436)
Individual Fixed Effects	Yes	Yes
N	149	261
R^2 (proj.)	0.3906	0.2814

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Standard errors clustered by individual

Table 13: Within-Person Returns to Manager-Directed Elaboration on Probability of Favorable Rating, Status Transitioners (Standardized)

Variable	**Women** **Model 21**	**Men** **Model 22**
Elaboration to Managers (std.)	-0.249***	0.061
	(0.0085)	(0.4105)
Manager	0.079	0.022
	(0.6868)	(0.8177)
Manager × Elab. to Managers	0.278*	-0.036
	(0.0650)	(0.6047)
Network Constraint (std.)	1.036*	0.365**
	(0.0548)	(0.0426)
Network Centrality (std.)	0.050	0.052
	(0.7444)	(0.5198)
Tenure (std.)	0.295*	0.190**
	(0.0801)	(0.0143)
Avg Rec Thread Depth (std.)	0.022	0.049
	(0.6452)	(0.6047)
Avg CC Thread Depth (std.)	0.119	0.216
	(0.7163)	(0.5573)
Count Emails Received (std.)	0.018	-0.033
	(0.8616)	(0.1649)
Count Emails Sent (std.)	-0.079*	0.005
	(0.0813)	(0.8058)
Unique Exchange Initiators	0.101	0.084
	(0.2886)	(0.1376)
Prop. Personal Topics Sent (std.)	0.800	-0.552
	(0.2075)	(0.4409)
Prop. Personal Topics Rec. (std.)	0.066	0.101
	(0.8525)	(0.7391)
Individual Fixed Effects	Yes	Yes
N	51	83
R^2 (proj.)	0.5972	0.3406

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Standard errors clustered by individual

Figures

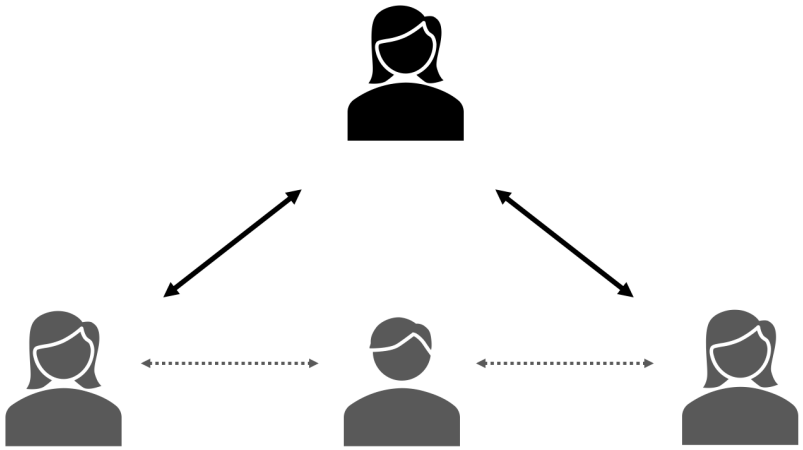

Figure 1: Cross-rank communication, because it requires greater contextualizing of problems and solutions, is more elaborated than within-rank communication.

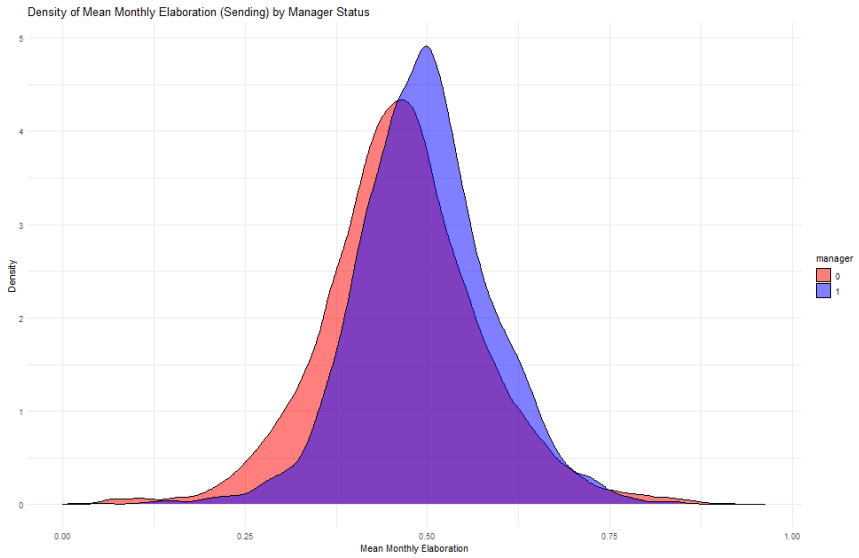

Figure 2: Density plot of average elaboration in emails sent by manager-status, aggregated to the monthly level.

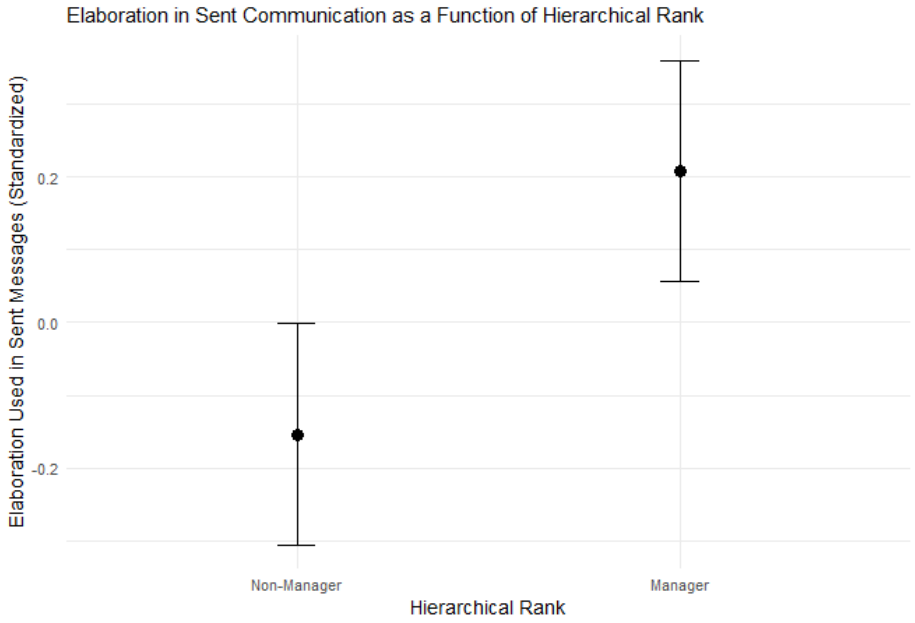

Figure 3: The estimated relationship between Hierarchical Rank and Linguistic Elaboration of work-related communication sent, based on linear regression with month-fixed effects and standard errors clustered by individual. Full controls for network position, department, and communication patterns (Model 6).

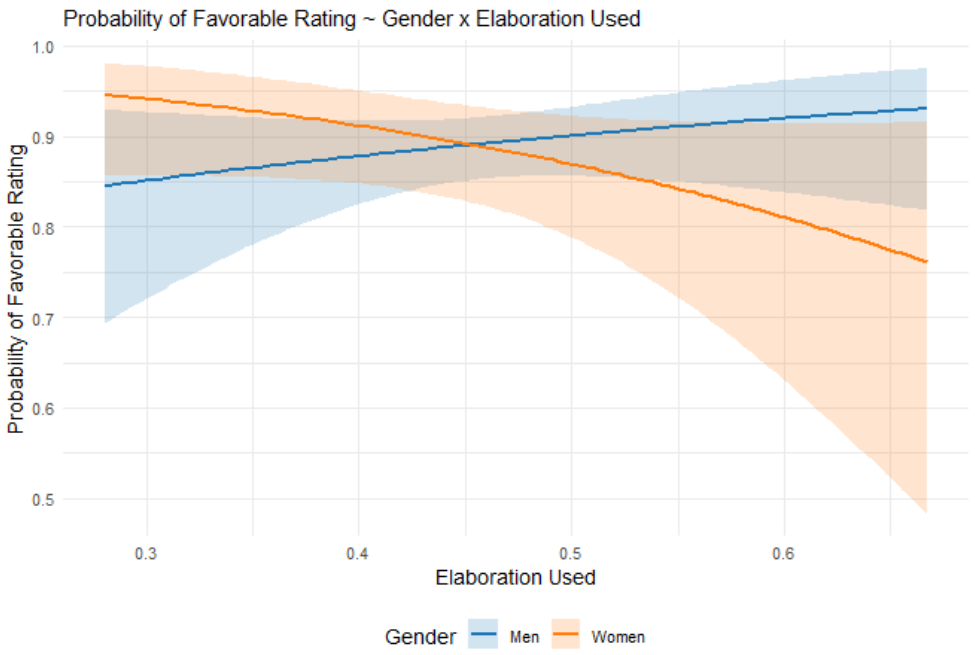

Figure 4: Probability of favorable performance rating as a function of Gender interacted with Linguistic Elaboration used. Predicted probabilities estimated via logistic regression for visualization; corresponding coefficient estimates reported in Table 8 use linear probability models with year-fixed effects and standard errors clustered by individual.

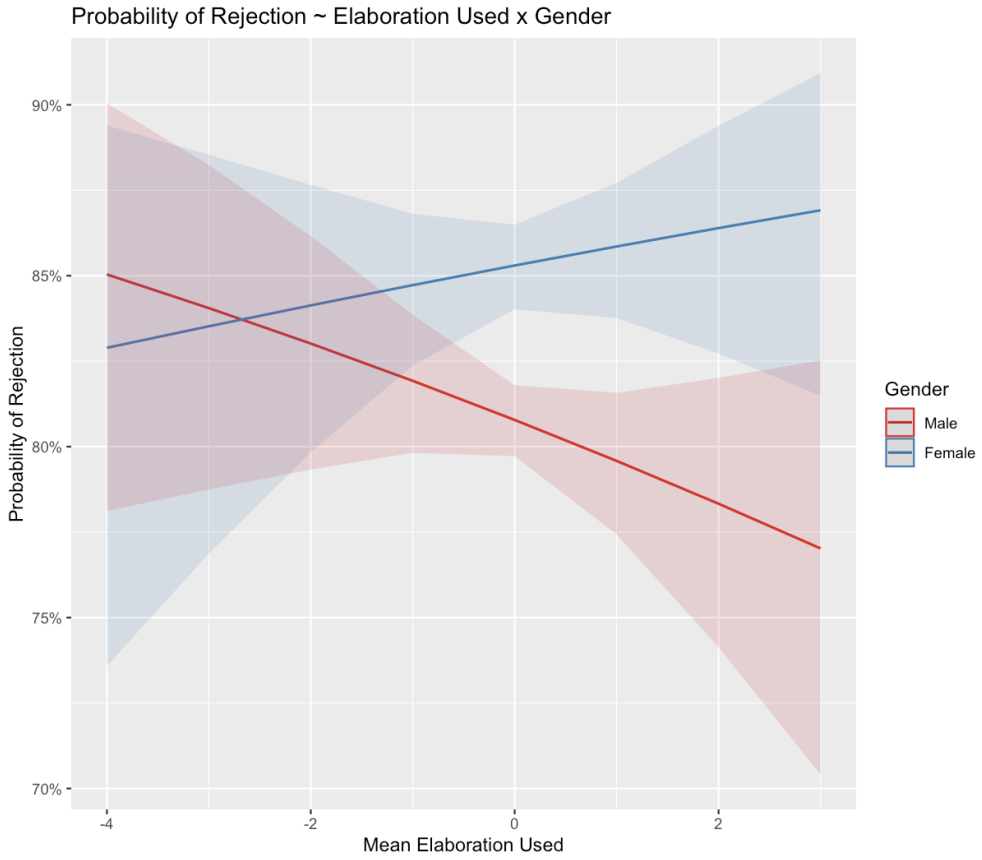

Figure 5: Probability of application rejection as a function of Gender interacted with Linguistic Elaboration used, based on GLM with a logistic link function, interviewer-fixed effects, topic-fixed effects, and heteroscedasticity consistent standard errors.

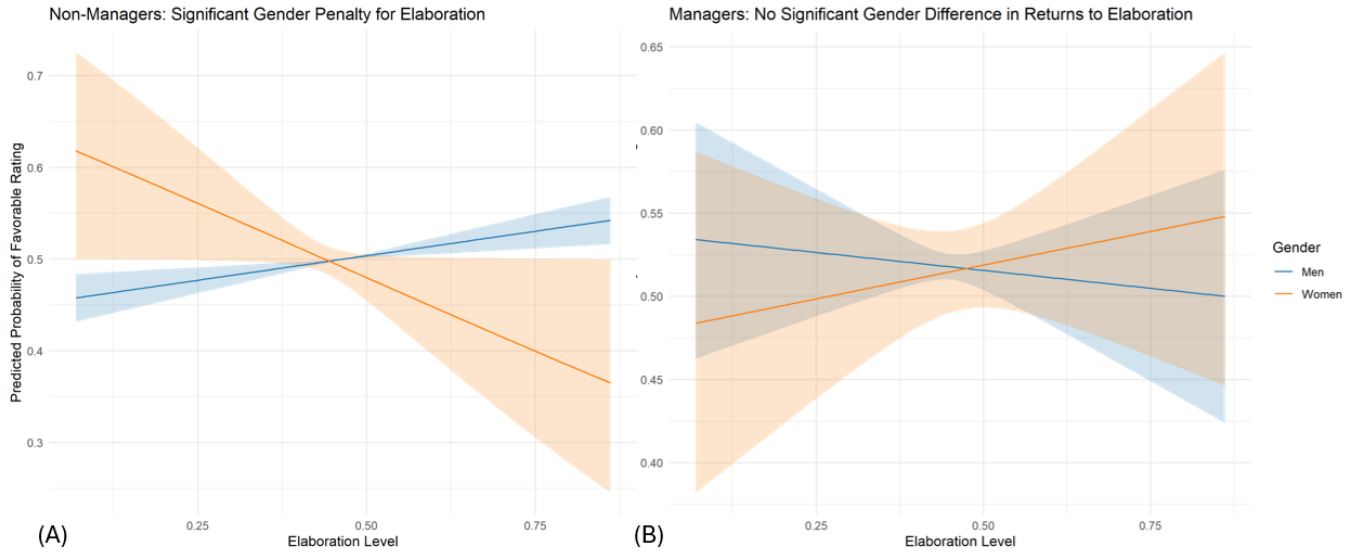

Figure 6: Probability of favorable performance rating as a function of Gender interacted with Linguistic Elaboration used, by manager status. Non-Managers shown in 6A and Managers shown in 6B. Predicted probabilities estimated via logistic regression for visualization; corresponding coefficient estimates reported in Table 10 use linear probability models with year-fixed effects and standard errors clustered by individual.

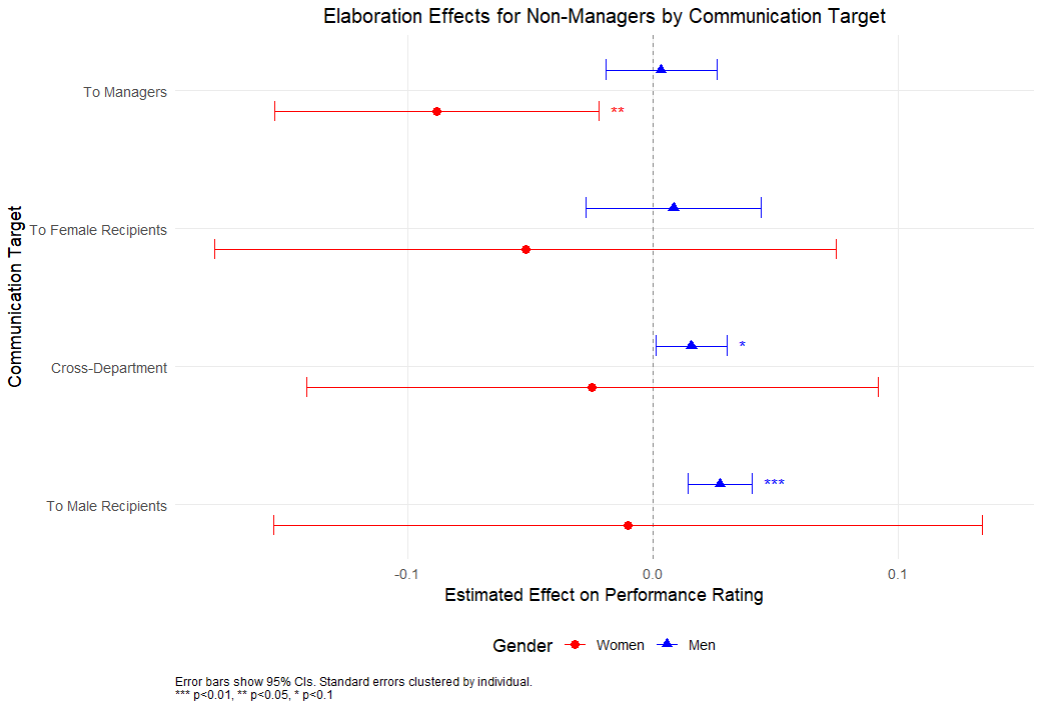

Figure 7: Estimated effects of elaboration on performance ratings for non-managers, decomposed by communication target, separate for men and women. Each estimate from a separate fixed-effects panel regression with controls for network position, communication patterns, and year effects. Error bars show 95% confidence intervals with standard errors clustered by individual.

Appendix A Alternative Explanations: Gender Differences in Problem-Related Communication

One potential explanation for the gender-based elaboration penalty in the email corpus is that women may be more prone to problems or mistakes in their work, forcing them to provide more context when explaining these issues. This would lead to both higher elaboration and lower performance evaluations. To test this alternative mechanism, I examine whether the elaboration penalty can be explained by differences in the extent to which men and women non-managers receive negative feedback, and the extent to which this feedback is differentially associated with probability of favorable performance ratings.

I apply the Valence Aware Dictionary and sEntiment Reasoner (VADER) to extract negative sentiment scores across all emails in the corpus (Hutto and Gilbert 2014). This measure produces a value between 0 and 1, indicating the proportion of the email's content that expresses negative sentiment. I aggregate this to the yearly level within employee by taking the mean negative sentiment of all emails received by the employee in a given year. This aggregation is meant to capture negative sentiment directed at the employee, which may reflect negative feedback, negative dispositions, or difficult situations.

Figure A1 depicts the density plots of average negative sentiment in emails received by gender, aggregated to the monthly level. There is considerable overlap in the distributions, however, women appear to receive slightly less negative emails on average than men. In Table A1, we explore this association more closely. Within the table, Models A1 and A2 show the results of the analysis re-estimated with controls for negative sentiment included. The coefficient for the interaction between gender and elaboration used is still significant. This provides evidence suggesting that the gender-based elaboration penalty is not driven by women responding to negative or critical communication.

More interesting, however, are the results for Model A3. Here we interact gender and mean annual negative sentiment received. Investigating the interaction effect more deeply probes this alternative explanation because it argues that women's elaboration may be primarily driven by having to do "damage control" or explain problems in response to criticism. By contrast, men's elaboration may be entirely different; it may reflect "leadership

behavior". This may explain why women non-managers appear to be punished for it whereas men non-managers appear to be rewarded for it.

I find a significant effect; however, it is in the opposite direction of what the alternative explanation would imply. Men experience greater performance evaluation penalties at higher levels of negative sentiment than women do. Figure A2 depicts this result more clearly. Here, we strikingly see that at higher levels of negative sentiment received, men tend to have lower probability of favorable rating received whereas women tend to have increased probability of favorable rating received. This pattern suggests that women's elaboration penalties are unlikely to stem from having to explain problems or mistakes. It may be that, for whatever reason, women navigate problem-focused communication more successfully than their male counterparts, or are more responsive to negative feedback in terms of performance.

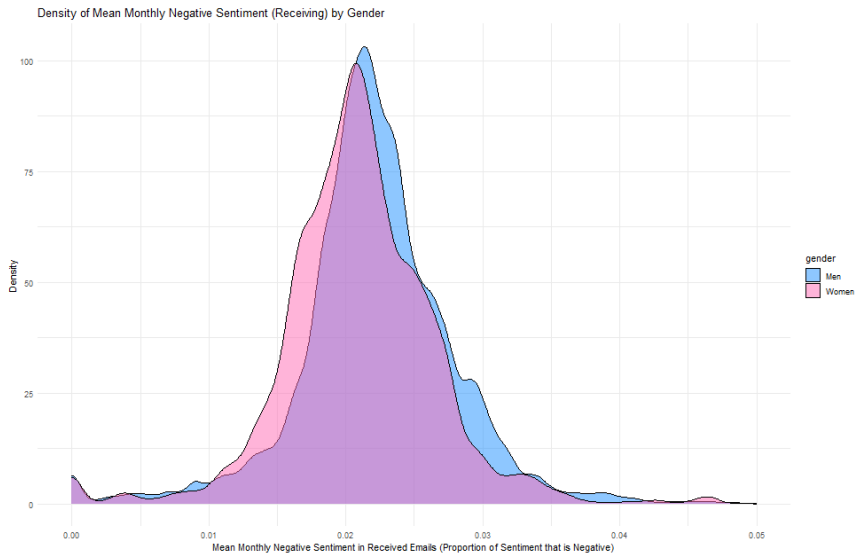

Figure A1: Density plot of average negative sentiment in emails received by gender, aggregated to the monthly level.

Table A1: Interaction between Gender, Linguistic Elaboration of Messages Sent, and Mean Negative Sentiment of Messages Received, on Probability of Favorable Rating

Variable	Model A1	Model A2	Model A3
Woman × Elaboration Used (std.)	-0.106**	-0.104***	
	(0.0427)	(0.0381)	
Woman × Mean Negative Sentiment (Rec.)			18.058**
			(7.4161)
Woman	-0.022	-0.015	-0.405**
	(0.0161)	(0.0269)	(0.1779)
Elaboration Used (std.)	0.032***	0.021*	-0.004
	(0.0111)	(0.0118)	(0.0186)
Mean Negative Sentiment (Rec.)	-2.786	-3.400	-7.455
	(4.0168)	(3.6900)	(5.3778)
Network Constraint (std.)	0.023*	0.015	0.020
	(0.0141)	(0.0516)	(0.0677)
Network Centrality (std.)	-0.011	-0.111	-0.108
	(0.0322)	(0.0834)	(0.0892)
Tenure (std.)	0.078***	0.067***	0.066***
	(0.0158)	(0.0109)	(0.0114)
Age (std.)	-0.051**	-0.050**	-0.048**
	(0.0199)	(0.0228)	(0.0218)
Sales	-0.062	-0.048	-0.050
	(0.0821)	(0.0828)	(0.0838)
Marketing	0.023	0.039	0.039
	(0.0417)	(0.0569)	(0.0522)
Tech	-0.037	-0.087	-0.087
	(0.0275)	(0.0531)	(0.0552)
Avg Rec Thread Depth (std.)	0.002	-0.001	0.002
	(0.0057)	(0.0052)	(0.0050)
Avg CC Thread Depth (std.)	0.011	0.024	0.028
	(0.0194)	(0.0254)	(0.0282)
Count Emails Sent (std.)		-0.007	-0.014
		(0.0130)	(0.0135)
Count Emails Received (std.)		0.034	0.034
		(0.0246)	(0.0263)
Unique Exchange Initiators		0.047	0.047
		(0.0472)	(0.0497)
Prop. Personal Topics Sent (std.)		0.017	0.017
		(0.0151)	(0.0162)
Prop. Personal Topics Rec. (std.)		0.007	0.007
		(0.0313)	(0.0354)
Period Fixed-Effects	Yes	Yes	Yes
N	5,593	5,191	5,191
R^2	0.0507	0.0627	0.0589
Adj. R^2	0.0477	0.0585	0.0547

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ Standard errors clustered by individual

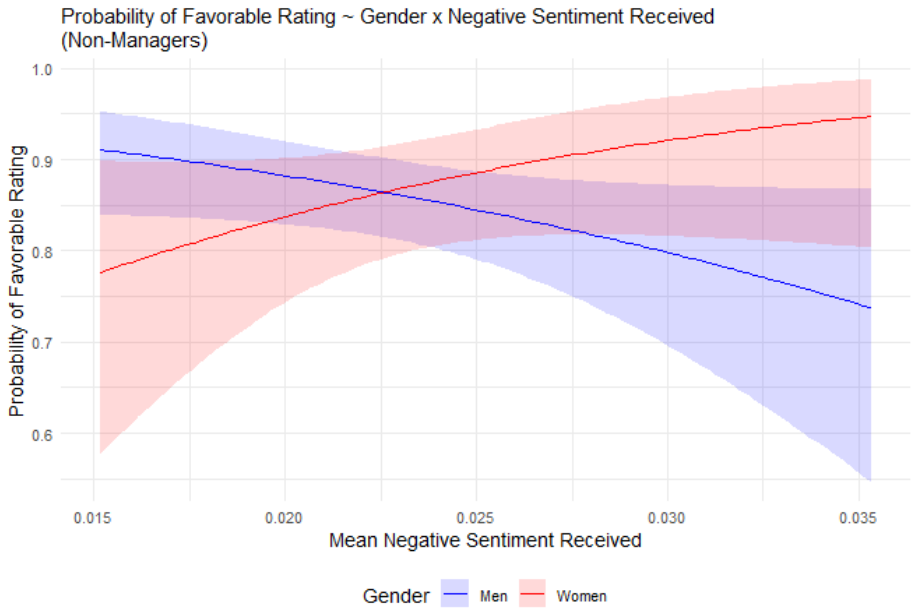

Figure A2: Probability of favorable performance rating as a function of Gender interacted with Mean Negative Sentiment Received in email communication, annually. Predicted probabilities estimated via logistic regression for visualization; corresponding coefficient estimates reported in Table A1 use linear probability models with year-fixed effects and standard errors clustered by individual.

Appendix B Alternative Explanations: Disentangling Status Violations from Performance Standards

A potential alternative explanation for my finding is that elaboration penalties reflect gendered sanctions for inefficient communication rather than gendered responses to authority displays. Under this view, divergent returns to elaborated language might simply reflect penalties for verbosity or wordiness, with women facing stronger penalties because they are held to stricter standards of communication efficiency. Thus, the mechanism for the observed disparity would align with gender-based performance expectations (Heilman 2001), where women's communication is scrutinized more heavily for signs of competence and professionalism, rather than gender-based status incongruity.

To evaluate this possibility, I calculate the mean word count of emails in the corpus, aggregated at the employee/year level. Given that word count distributions are typically right-skewed (Sichel 1975; Piantadosi 2014), I log-transform word counts before estimating their relationships with other variables in the sample. Figure B1 illustrates the correlation between linguistic elaboration and word count in the email dataset. Here, word count is derived by first log-transforming the yearly average number of words per email and then averaging these log-transformed values for each employee. Similarly, linguistic elaboration scores are standardized and aggregated at the employee level. The results show that elaboration is moderately correlated with verbosity ($r = 0.27$, $p < 0.001$, 95% CI: [0.27, 0.36]), but this emphasizes that elaboration captures distinct linguistic qualities beyond mere message length.

To further address this alternative explanation, I conduct a series of regression analyses to examine the key associations underpinning my theoretical argument, now including controls for word count. Model B1 first establishes that while verbosity strongly predicts elaboration ($\beta = 0.912$, p<0.001), it does not fully explain elaboration patterns: both Woman ($\beta = 0.065$, p<0.05) and manager ($\beta = 0.163$, p<0.001) remain significant predictors after controlling for word count. This further underscores that elaboration captures distinct linguistic qualities beyond verbosity – qualities that systematically vary with both gender and hierarchical position.

Models B2 and B3 deliver key tests of the competing mechanisms. If elaboration penalties reflected efficiency standards, we would expect word count to show negative effects or at least mitigate elaboration's relationship with performance. Instead, in both models, word count shows positive associations with performance ratings (β = 0.089, p<0.01 for word count), while women face specific penalties for elaboration use (Woman × Elaboration: β = -0.076, p<0.05). This pattern is difficult to reconcile with an efficiency-based explanation but aligns perfectly with status violation dynamics.

Model B4 provides even stronger evidence against the performance standards account. The interaction between gender and word count (β = -0.112, p=0.15) is not significant, indicating women do not face systematically harsher penalties for verbose communication. Moreover, the consistently positive main effect of word count on performance ratings (β = 0.151, p<0.001) directly contradicts the notion that organizational evaluators penalize inefficient communication, or respond much differently to it based on gender.

Together, these results strongly favor the status violation mechanism over divergent performance standards. Women's elaboration use predicts lower performance ratings not because evaluators scrutinize their communication efficiency more harshly, but because elaboration has become associated with managerial authority in this organizational context. The penalty thus reflects status violation rather than gender incongruence: women are sanctioned for authority-signaling behavior even when that behavior is stereotypically feminine.

Table B1: Regression Results for Elaboration, Gender, and Workplace Outcomes

Variable	Model B1: Elaboration Usage	Model B2: Performance Rating	Model B3: Performance Rating	Model B4: Performance Rating
Woman × Elaboration		-0.076**	-0.074**	
		(0.0350)	(0.0300)	
Woman × Log Word Count				-0.112
				(0.0770)
Elaboration		0.011*	-0.003	-0.025*
		(0.0070)	(0.0120)	(0.0150)
Woman	0.065***	-0.002	0.001	0.378
	(0.0240)	(0.0110)	(0.0160)	(0.2630)
Log Word Count	0.912***	0.089***	0.111***	0.151***
	(0.0860)	(0.0140)	(0.0220)	(0.0310)
Network Constraint	0.045	0.044**	0.044	0.043
	(0.0660)	(0.0190)	(0.0420)	(0.0480)
Centrality	-0.010	0.012	-0.064	-0.058
	(0.0750)	(0.0260)	(0.0690)	(0.0710)
Tenure	0.018	0.043***	0.033***	0.031***
	(0.0270)	(0.0090)	(0.0090)	(0.0080)
Age	0.033	-0.038**	-0.036*	-0.035*
	(0.0200)	(0.0180)	(0.0190)	(0.0200)
Manager	0.163***	0.068***	0.055***	0.050**
	(0.0200)	(0.0180)	(0.0200)	(0.0200)
Sales	-0.393***	-0.077	-0.068	-0.069
	(0.1010)	(0.0710)	(0.0670)	(0.0670)
Marketing	-0.106	0.013	-0.003	0.000
	(0.1150)	(0.0350)	(0.0480)	(0.0480)
Tech	-0.073	-0.027	-0.065	-0.067
	(0.1480)	(0.0240)	(0.0490)	(0.0440)
Conversation Depth (Rec)	-0.015	0.004	0.005	0.007
	(0.0190)	(0.0050)	(0.0050)	(0.0040)
Conversation Depth (CC)	0.027	-0.003	0.019	0.019
	(0.0410)	(0.0310)	(0.0320)	(0.0300)
Count Emails Sent	0.144***		0.005	0.003
	(0.0100)		(0.0140)	(0.0160)
Count Emails Received	-0.014		0.019**	0.022*
	(0.0480)		(0.0100)	(0.0120)
Prop. Personal Emails Sent	-0.075***		0.006	0.006
	(0.0190)		(0.0160)	(0.0160)
Prop. Personal Emails Received	-0.046		-0.001	-0.001
	(0.0460)		(0.0260)	(0.0260)
N	11,503	6,584	6,173	6,173
R^2	0.337	0.051	0.057	0.054
Adj. R^2	0.336	0.048	0.054	0.050

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Standard errors clustered by individual

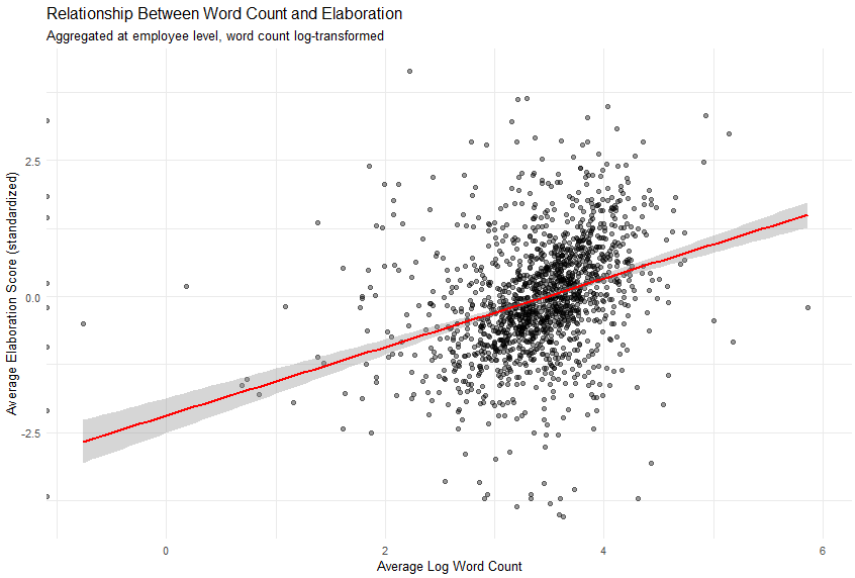

Figure B1: Scatterplot showing the relationship between log-transformed average word count and standardized average elaboration scores, aggregated at the employee level. A positive linear trend is observed, indicated by the red regression line.

Appendix C SEMEL: Stacked Ensemble Model of Elaborated Language

Overview of SEMEL

The Stacked Ensemble Model of Elaborated Language (SEMEL) was developed to quantify the level of linguistic elaboration in natural language text. Rooted in Bernstein (1971)'s sociolinguistic theory of elaborated versus compact language codes, SEMEL aims to determine whether a given piece of text is context-dependent (compact) or explicitly detailed (elaborated). This approach is significant in understanding the underlying social dynamics between interlocutors, such as formality, intimacy, and shared cultural knowledge. SEMEL employs an ensemble of sub-models to systematically analyze linguistic features, combining them to produce a quantifiable elaboration score.

Methodology

Stacked Ensemble Design

SEMEL's architecture relies on five core learning assignments: part-of-speech (POS) tagging, syntactic parsing, sentence embedding, co-reference resolution, and neural classification. These assignments are processed independently, and their outputs are aggregated to predict elaboration scores. The ensemble is constructed via a stacking mechanism, where fully trained sub-models solve individual linguistic tasks, and a final neural classifier combines the outputs to evaluate the level of elaboration in a block of text.

Feature Extraction and Processing

SEMEL utilizes two forms of syntactic parsing – dependency and constituency parsing – to identify hierarchical relationships within text. Dependency parsing determines the grammatical relations between words, while constituency parsing identifies the structural units within sentences, such as noun and verb phrases. This dual parsing approach enables SEMEL to distinguish between compact and elaborated syntax by capturing both

the grammatical relationships and the constituent structures present in each text block (Jurafsky and Martin 2020).

Additionally, SEMEL generates word embeddings using the fast-Text model, a method that creates vector representations of words based on their character n-grams (Bojanowski et al. 2016). These embeddings help quantify semantic similarity and contribute to identifying elaborated language, as elaborated texts tend to exhibit greater lexical diversity and descriptiveness.

The co-reference resolution component identifies clusters of referents within a text to evaluate the explicitness of references. Compact language often relies on implicit or pronoun-heavy referencing, whereas elaborated text makes more frequent use of explicit referents (such as proper nouns). SEMEL leverages NeuralCoref 4.0 to quantify these relationships, capturing the extent to which pronouns or other reference chains are used (Wolf 2017).

In addition to outputs from sub-models, SEMEL computes several engineered features, such as the F-score for formality (Heylighen and Dewaele 1999), the Coleman-Liau readability index (Coleman and Liau 1975), type-to-token ratio (TTR), and measures of term rarity based on the Brown (Francis and Kucera 1979) and Reuters (Lewis, Yang, Rose, and Li 2004) corpora. These features allow for a more nuanced quantification of the text's elaboration.

Training Procedure

The training data consists of modified excerpts from the Spoken category of The Corpus of Contemporary American English (CoCA) (Davies 2008) and The Corpus of Linguistic Acceptability (CoLA) (Warstadt, Singh, and Bowman 2018). Text blocks are hand-labeled with elaboration scores ranging from 0 to 100, depending on the degree of explicitness and descriptiveness. SEMEL is trained using a pairwise classification approach, where text block encodings are paired, and the neural classifier learns to identify the more elaborated block. A weighted cross-entropy loss function is employed, with penalty weights proportional to the difference in elaboration scores between paired blocks, ensuring the classifier accounts for the granularity in elaboration levels.

The classifier's training relies on a combination of hand-labeled examples and machine-augmented blocks generated via POS-tag substitution, resulting in a diverse set of training data. Pairwise comparisons enable the model to derive a list-wise ranking of elaboration scores by evaluating text blocks relative to each other. The final elaboration score is computed by comparing each input block to a reference set of blocks, which ensures consistency in evaluation.

Evaluation and Performance

Performance on Validation Set

SEMEL's performance was evaluated using pairwise classification accuracy and list-wise Spearman rank-order correlation. Final-epoch pairwise classification accuracy reached 88.7%, while the rank-order correlation was 0.964, indicating a high level of reliability in differentiating text elaboration. Ablation studies further demonstrated the relative importance of the model components, with dependency parsing and co-reference resolution proving critical for accurate elaboration quantification. The results suggest that syntactic features, explicit referent use, and lexical diversity are key drivers in distinguishing between elaborated and compact codes.

External Validation on the Wizard of Wikipedia (WoW) Dataset

To evaluate SEMEL's construct validity, it was deployed on the Wizard of Wikipedia (WoW) dataset (Dinan et al. 2019). The WoW dataset consists of dialogues where one participant (the "wizard") provides detailed, elaborated responses on various topics, while the other participant (the "apprentice") asks questions or makes comments. The aim was to determine if SEMEL could consistently attribute higher elaboration scores to the wizard, as this role was designed to incorporate greater context and provide more detailed information, except in a highly natural way. Similarly, the apprentice's role was to seek information about the topic, in a natural and non-obvious way.

Results from the WoW validation showed that SEMEL successfully identified a statistically significant difference in elaboration levels between the wizard and apprentice roles.

Specifically, the wizard's elaboration score was consistently higher, which aligns with the role's purpose of integrating external knowledge into the conversation. This outcome serves as an informal measure of SEMEL's construct validity, demonstrating that it effectively captures elaborative content that stems from greater contextual knowledge, beyond superficial features of text like sentence length or position in dialogue sequence.

The figures below provide insight into SEMEL's methodology and its evaluation process, and demonstrate the systematic flow from initial input processing to final elaboration scoring.

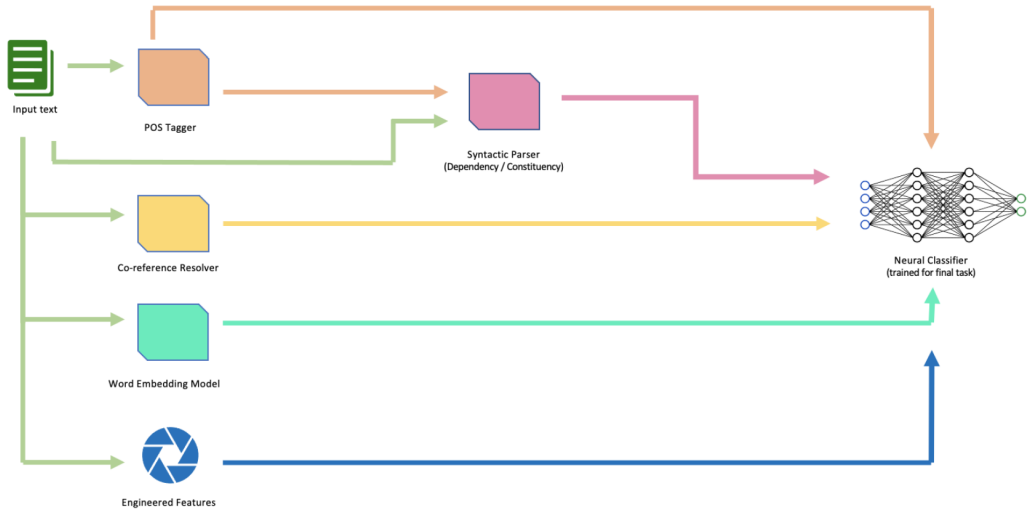

Figure C1: This figure illustrates the process flow of encoding input text blocks, starting with initial parsing, through feature extraction, to final classification.

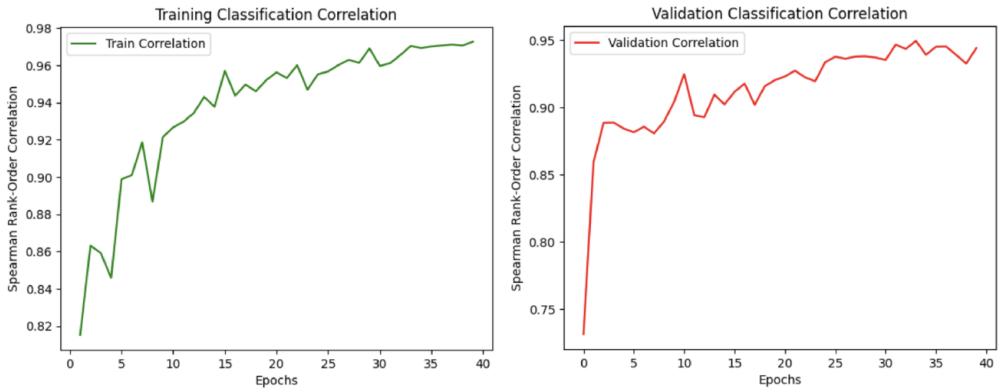

Figure C2: This figure shows the progression of list-wise correlation throughout training, highlighting SEMEL's increasing ability to rank text blocks by elaboration level

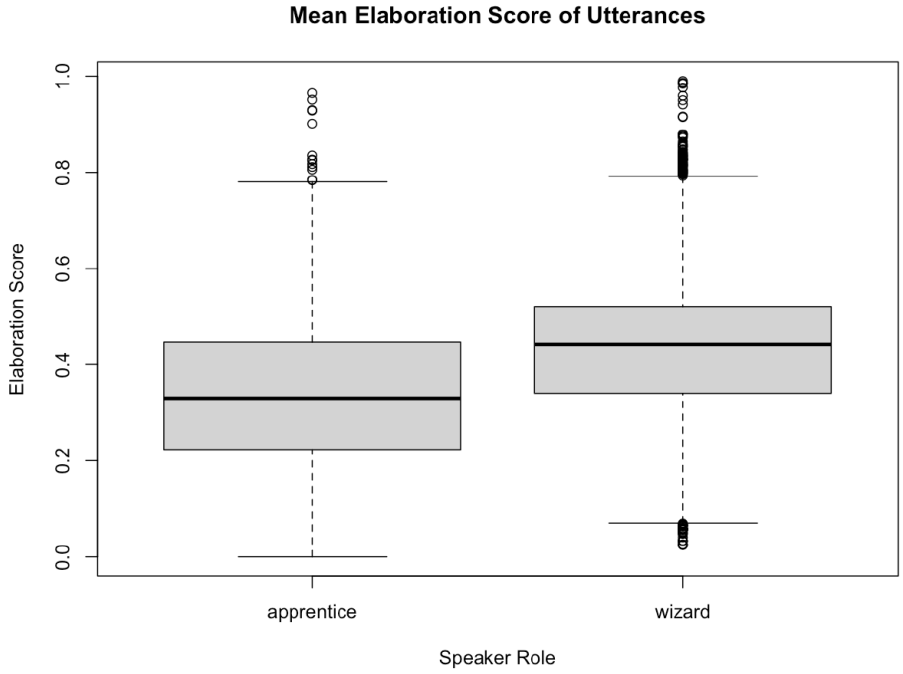

Figure C3: Unconditional Mean Elaboration Score of Utterances, by Speaker Role in the WoW dataset (N = 113, 853).

Appendix D Additional Illustrative Examples of Linguistic Elaboration

This appendix provides a fuller set of illustrative examples underlying the discussion in the Measuring Linguistic Elaboration subsection. Whereas Table 3 in the main text presents abbreviated excerpts selected to make the contrast between higher and lower elaboration legible, Table D1 reproduces those same examples in fuller form. The purpose is expository rather than inferential: the examples are included to make the construct more concrete, not to suggest that elaboration maps perfectly onto managerial status or any other sender characteristic. Both managers and non-managers produce communication spanning the full range of elaboration in the data. All examples are initial emails in a thread, which helps hold constant differences that might otherwise arise from message position within an ongoing exchange. The examples are lightly edited for readability and anonymized, with placeholder indices resetting within each email excerpt.

Table D1: Fuller Illustrative Examples of Higher and Lower Linguistic Elaboration

Example	Sender status	Thread status	SEMEL	Excerpt	Key feature
Panel A. Higher-elaboration examples					
H1	Manager	Initial	.901	"[NAME1] reached out to discuss ways she could contribute to the services program. Her initiative was welcome given the high expectations set by management and the short runway for results. At present, the team does not have a full-time member representing the voice of the customer, and several initiatives in the pipeline would benefit from someone prepared to take the lead under a design-thinking framework. I suggested that [NAME1] shadow the team over the next couple of weeks so that she can see the process in action and better align her interests with the team's needs. This should require no more than a limited time commitment initially, and we can regroup in early February to discuss next steps."	Actors identified; background, gap, and next steps made explicit.

Continued on next page

Example	Sender status	Thread status	SEMEL	Excerpt	Key feature
H2	Manager	Initial	.901	"I spoke with [NAME1], and we will only need to supply public IPs for [LOCATION1]; the remaining locations are out of scope. Any remote recruiting or HR personnel will need to VPN back to [LOCATION1] to access the site. The full document is attached, but I have also pasted the relevant section below. Customers using credit-reporting services must register authorized public static IP addresses for an additional authentication check. Internal private DHCP addresses are not relevant for this requirement. Please provide the list of public static IP addresses that should be authorized."	Scope clarified; policy requirement, exclusions, and request all made explicit.
H3	Manager	Initial	.889	"What [NAME1] said is accurate. The phones are an independent brand and are compatible with multiple backend systems; in fact, they are the same phones we use in [LOCATION1]. The switch providing power to the phones is also the same brand and model. The problem appears to be local to the temporary space. Humidity, wide temperature swings, and voltage drops caused by daisy-chained computer systems could all be contributing factors. [ORG1] is replacing each phone that burns out, and all of these units were brand new. We are now awaiting new power supplies and battery systems in an effort to troubleshoot the issue further."	Problem framed with background, possible causes, and prior troubleshooting.
H4	Non-manager	Initial	.760	"I had actually been meaning to touch base with you about this one. What was said to the customer was accurate: I did mention that at some point we would have the option to use a new [ORG1] product that relies on cellular rather than internet connectivity. That said, I have not heard that the option is available right now, and the customer still does not have a stable connection of his own. I am not sure how best to proceed. We could install the current unit and wait for him to provide a router, but I will also check with our post-install team to see what they recommend."	Background, prior communication, current constraint, and possible next steps made explicit.

Continued on next page

Example	Sender status	Thread status	SEMEL	Excerpt	Key feature
H5	Non-manager	Initial	.760	"We are running into technical issues updating the templates. For the time being, the templates and the string sizer will therefore disagree about which inverter is correct to use. The string sizer document is always correct. We are no longer using [TECH1], and in many cases we are using [TECH2] instead of the [TECH3] or [TECH4], even for a single face. Almost all [TECH5] systems have now been replaced by [TECH6] systems. If you are doing QA, please fail people consistently for using the wrong inverter if the system was sized after this email went out."	Explains temporary inconsistency, states governing rule, and clarifies implications for action.
Panel B. Lower-elaboration examples					
L1	Non-manager	Initial	.440	"That makes sense, and I feel like I did that and left no stone unturned. Would you mind listening to the entire call before giving me feedback suggesting that I may not have done so? I feel like I pushed that point as far as I could, but not every customer responds to it. I also think that is one of my stronger areas—selling on other values and encouraging them to look ahead. Perhaps we can work on ways for me to do that even more effectively, using this call as one example."	Immediate exchange emphasized; less explicit broader organizational scaffolding.
L2	Non-manager	Initial	.413	"This project had some complicated inspection call-outs by the AHJ. We need to add an AC disconnect at the main house, which is separate from the workshop housing the PV modules. It took some time for the team to work through the design-change options and coordinate with the customer. The work requires a substantial amount of wiring, but it eliminates the need for trenching between the two buildings. The work is scheduled to be completed today, with inspection either today or tomorrow."	Issue explained more linearly and directly, with less layered contextual framing.
L3	Non-manager	Initial	.447	"From the photos, it does not look like the wire is stranded. It appears to be THWN or THHN, and it may be either [TECH1] or [TECH2]. If it is [TECH1], it is out of compliance, but if it is [TECH2], we are fine. Based on the information available, this should be acceptable. I will revise the plan set to reflect the new information."	Narrow technical issue resolved with limited broader situational context.

Continued on next page

Example	Sender status	Thread status	SEMEL	Excerpt	Key feature
L4	Manager	Initial	.437	"When I used to manage the account, I would run a report to see who was not using the product and tell them I was deleting it due to inactivity. I would just confirm that the people you want to delete definitely do not need it. Is there anything you need from me? I support whatever you decide. My only opinion is that it is worth warning people before deleting, but I am in favor of saving the money."	Advice given directly, with relatively little broader contextual scaffolding.
L5	Manager	Initial	.449	"What is going on here? [NAME1] reached out to let me know that [NAME2] is unhappy with the experience so far and wants to cancel because the call did not come at the right time and then was still perceived as poor once it did happen. These leads are now either cancelling or saying they will cancel. What happened, and is there a way to save these? I am looking for insight here."	Immediate problem and request foregrounded, with less extended contextual development.