

Write Here, Write Now!: An Experimental Study of Group Maintenance in Collaborative Writing

Jeremy Birnholtz
Communication Studies Dept.
Northwestern University
Evanston IL 60208
jeremyb@northwestern.edu

Stephanie B. Steinhardt
Communication Department
Cornell University
Ithaca, NY 14853
sbg94@cornell.edu

Antonella Pavese
Google, Inc.
76 9th Ave
New York, NY 10011
antonellap@google.com

ABSTRACT

Writing documents together using collaborative editing tools has become extremely common with the widespread availability of tools such as Google Docs. The design of such tools, rooted in early CSCW research, has historically been focused on providing awareness of the presence and activities of one's collaborators. Evidence from a recent qualitative study, however, suggests that people are also concerned about how their behaviors – and they themselves – will be perceived by others; and take steps to mitigate possible negative perceptions. We present an experimental study of dyads composing documents together, focusing in particular on group maintenance, impression management and relationship-focused behavior. Results suggest that communication is positively related to social relations, but only for synchronous writing in a shared space; the reverse can be true in asynchronous commenting and editing.

ACM Classification Keywords

H.5.3 Group and Organization Interfaces

Author Keywords

Collaborative writing; awareness; group maintenance.

INTRODUCTION

Writing documents together has long been a common occurrence in both organizational and educational contexts. Early CSCW research recognized the potential of tools to support these activities and experimented with several such tools (e.g. ShrEdit [24], Quilt [17], SASSE [26]), but features and ideas explored in these systems (and others) have only recently gained mainstream appeal via common tools such as Google Docs and the Microsoft Office suite of applications. As such, collaborative writing tools have not been a significant focus of recent CSCW research.

Where recent CSCW research has focused on systems for collaborative writing, it has largely focused on technical problems, such as updating revision histories [15], or built on early work focusing on the importance of providing awareness information (e.g., [8]). Awareness of who else is

present and what they are doing in a document can help people understand changes and coordinate their efforts, and to answer questions about how things got the way they are (e.g., [13, 28]). At the same time, however, awareness information on its own is often insufficient to support the collaborative writing process.

Consider an academic paper being written by Chris and Alex. Alex writes the first draft of the literature review, and Chris agrees to provide some edits and comments on this. Chris's edits are tracked by the revision history feature in the word processor, so Alex can see exactly what has been changed and when it was changed. Chris also makes a few comments on parts he thinks could be changed. When Alex looks at the edited version, he is surprised to see how many changes Chris has made; and that many of the changes are little, picky changes to things that Alex felt were important to the character of the document. Alex is upset with Chris and takes pains to carefully explain why he had written the document the way he did.

While this scenario is fictional, it is based on a combination of real situations described by Birnholtz & Ibara [4]. It highlights that others' changes to a document can have negative relational effects, and that today's collaborative writing software is often not geared toward a relational approach to collaborative writing. There is reason to believe, however, that such a perspective could be useful. In their study, Birnholtz & Ibara showed that people are conscious not just of interpreting the information provided by awareness features (e.g., who made a particular edit), but also about how others will interpret this information (e.g., what will my co-author think of me if I make this edit?).

In this way, editing and work in shared documents can be considered performative behavior, in that it is executed in the presence of others and with an eye toward how they will interpret it. This is consistent with Goffman's theory of impression management [12], and theories of group behavior that highlight group maintenance as a key component of group activities [19].

In this paper, we describe a lab experiment exploring the use of communication and awareness features by dyads writing documents together. Results suggest that communication was useful, but had a positive relational impact primarily in synchronous writing; and that text communication complemented, rather than substituted for, editing.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2013, April 27–May 2, 2013, Paris, France.

Copyright © 2013 ACM 978-1-4503-1899-0/13/04...\$15.00.

BACKGROUND

Awareness & Collaborative Writing

While collaborative writing tools have only recently become popular beyond specialized domains [23], they were an early topic of research in the CSCW community. The earliest systems (e.g., Quilt [17], ShrEdit [24]; Grove [9]) were designed to experimentally explore the utility and function of user roles, permissions and access controls.

Through studies of these and subsequent systems, it became clear that a key issue in collaborative writing is awareness [8]. Specifically, awareness information that indicates who else is present in a space and what they are doing is vital because it facilitates knowing what is taking or has taken place within a document, how certain elements of the document may have changed since it was last viewed, and when changes were made [28]. These features enable users to better understand and coordinate each other's efforts. Through laboratory studies of collaborative writing using these systems (e.g., [17, 24, 26]) and field studies with those who wrote documents together (e.g., [2], [25]), researchers examined elements of the writing process and the utility of specific features in supporting identified tasks.

These examinations of the writing process made it clear that writing documents together often involves many different modes of collaboration and coordination, ranging from a single writer doing most of the work, but seeking comments from a collaborator; to joint composition of text by multiple users in real time [25]. Indeed, some elements of shared writing tasks are often completed synchronously or nearly so (e.g., joint composition of text, discussion of content), while others occur asynchronously (e.g., commenting on a collaborator's early work, editing and proofreading, etc.). Commenting and providing feedback to others is a key component of all of these processes [20].

In all of these components of collaborative writing, awareness features provide information about others' behavior that can aid coordination [22]. In particular, awareness information can answer questions about who else is present or has access to the document, what others are doing, and what has changed since the document was last viewed. Some systems also make draft text invisible to others until it is "committed" or "shared" by its author (e.g., [20]), whereas others immediately share updates in real time (e.g., [24]).

Moreover, awareness information can play multiple roles in coordination. At a basic level, it draws attention to specific behaviors of others, as when differences are highlighted [21]. Awareness information can also play a more subtle role, by providing cues that together allow users to understand each others' intentions and roles in the writing process. Dourish & Bellotti [8] suggest that increased awareness of other users and one's own contributions allows collaborators to understand and situate their work.

In addition to features that provide awareness information about others within a document, another source of such

information is communication between participants. Posner et al. [25] discuss the importance of communication in assigning tasks and coordinating effort. Erkens et al. [10] experimented with a tool that allowed communication in planning shared effort. Neuwirth et al. [20], Kim and Eklundh [16], and Bietz [3] all discuss the provision of feedback and how this is affected by both media and co-authors' relationships with each other. Birnholtz and Ibara, moreover, point out that communication can be used to explain potentially threatening behavior and avoid possible negative relational effects [4].

Awareness and Social Relationships

As noted above, awareness features in collaborative writing systems provide information about others' behavior in documents. This information, however, often excludes contextual details or explanations that could affect perceptions of behavior. There are several reasons to believe that these perceptions matter, however.

First, theories of group behavior point to the importance not just of a group's ability to complete a task at hand, but also to ensure, through relationship and group maintenance, that the group functions smoothly and is able to work together on subsequent tasks [7, 14, 19]. When people perceive others to be invading their space or negatively affecting a shared document, this can affect not only the document, but also the relationship between collaborators, people's perceptions of who owns and is responsible for the document [5] and their ability to work together.

In this way, working in a shared document is quasi-public behavior in that it is visible to others. Those others, moreover, may perceive it in ways that have relational consequences. This reflects Goffman's [11,12] theories of self-presentation and identity management, which use a dramaturgical metaphor to argue that people carefully consider what details of themselves to reveal "on stage" and what to hide "backstage." People are concerned not only about the details of their behavior but also about how it is perceived by others, and how these perceptions might affect how individuals themselves are perceived.

Consistent with these ideas, Birnholtz and Ibara [4] described instances in which collaborative writers intentionally avoided behavior that might be perceived negatively, or used text explanations to mitigate possible relational damage. Being conscious of multiple possible interpretations of one's actions is consistent with literature on accounting for behavior [6, 27], which document the use of language to explain potential face or relationship threats.

Despite the acknowledged importance of group maintenance in collaboration, however, the relationship between awareness features, communication and social relationships have received scant attention in literature on collaborative writing. Most work on writing systems has focused on providing awareness of others' behavior (see above). Neuwirth and colleagues [20] did, however, show that the modality for feedback presentation can affect its perception.

Using systems that are common today, there are several elements of writing behavior that we might expect to be affected by concerns about self-presentation, and situations in which communication between participants could be used to explain potentially face-threatening behavior, such as editing a collaborator's work, or to make suggestions instead of editing. As such, we would expect that people who have access to communication and awareness features to behave differently than those who do not.

THE PRESENT STUDY

To explore these issues, we ran a laboratory study in which dyads used Google Docs to compose individual documents, asynchronously edit and provide feedback on their partner's work, and then synchronously synthesize their two documents into a single summary document. In one condition, participants were given access to communication and revision history features; in the other, they were not.

Research Questions

We set out to answer three key research questions. Given the prevailing focus in the literature on awareness of others' behavior and the use of communication primarily for coordination purposes, we were first interested in understanding the tools and techniques people use to communicate, as well as what they talk about. We asked:

RQ1a: Do people communicate when writing together? If so, what channels do they use to communicate?

RQ1b: Is there a relationship between communication and writers' social relationships? Does communication appear to serve a relational purpose in collaborative writing?

Second, we were interested in the relationship between editing and social relationships. An awareness-based approach would suggest that people with changes to make will edit directly, and this should not affect social relationships. A relational approach would suggest that people will make suggestions, but edit less when communication tools are available. This should positively affect social relationships between participants. We asked:

RQ2a: Is there a relationship between editing and social relationships?

RQ2b: What is the relationship between editing and communication?

Finally, we wondered if these effects were different for asynchronous and synchronous phases of document creation. We therefore explored RQs 1 and 2 in a task that involved both asynchronous and synchronous writing.

Design

The design of our study was based on a within-subjects design. Pairs of strangers completed two different writing tasks with and without communication and awareness features available. Availability of these features was manipulated between participants, and the type of writing (asynchronous vs. synchronous) was manipulated within. The task was intended to mimic the stages of collaborative writ-

ing documented in studies cited above. It involved two phases:

Asynchronous Phase: The first task phase involved the asynchronous creation and editing of text. Each participant was instructed to individually write a 250-word statement taking one of two possible positions on a controversial issue, with each person assigned one side. This document was written in Google Docs and was then shared with their assigned partner, who was told to provide feedback and/or edit for strength of argument, grammar, and document flow.

Synchronous Phase: The next phase involved synchronous composition and editing. Participants were told to generate a new 250-word document that summarized both of their assigned positions on the issue. They were told that this document would be reviewed for clarity and correctness, and to ensure that it met the assigned parameters. Those who met these criteria were awarded an extra \$5 each.

In one condition, which we call "communication features" (CF), participants were given a chat box for real-time text interaction, the capacity to insert comments linked to specific text, and the ability to view a revision history showing what had changed in a document and who changed it. In the other condition, ("no communication features" or NCF, participants were told not to use these features.

As the two phases involved qualitatively different tasks, we do not compare them statistically in presenting our results.

Participants

There were 130 total participants (65 dyads) in the study. Thirty had to be removed from the data set for failure to follow instructions, either because they used features they were instructed not to use (1 case) or because they did not understand the task. This left a final data set of 100 (50 dyads). The study took place at a large US university from October - December, 2011. Participants were recruited individually (i.e., not in pairs) using a campus-wide web-based system, and were compensated with either cash (\$10) or class extra credit. Mean age was 20.33 (SD = 1.73), and 41 were female (59 male). Dyads were assigned at random to either the CF or NCF condition. Assignment to these conditions was initially nearly equal, but the removed cases meant that the final count was 20 NCF dyads and 30 CF. Most participants (79%) reported some prior experience with collaborative writing software.

Task and Materials

The task was designed to incorporate elements of both asynchronous and synchronous modes of writing together. Participants were initially presented with a description of the issue they were to write about, along with a set of short bullet points describing their assigned position on the issue.

The issue selected was a real 2011 controversy in which a captive primate had "stolen" a photographer's camera, and taken several photographs. As the story and corresponding photographs made its rounds in the news, an agency acting on behalf of the camera owner claimed copyright infringe-

ment and requested all photos be taken down off all sites. However, the service’s right to do so was challenged by those who believed, consistent with copyright law, that the images belonged to the content creator (i.e., the primate). One participant was assigned the position that the human photographer owned the photos, and the other was assigned the position that the images were in the public domain, as the creator (the primate) cannot own them. We selected this issue because it was real but had no clear correct answer, and because participants’ views were unlikely to be colored by their prior experience or political stance.

Participants were seated in separate rooms in our laboratory, each of which contained a PC and a 17” LCD monitor. Writing tasks were completed using Google Docs, via accounts created for the experiment. Three new documents were created for each dyad. Two contained the initial description and bullet points for the individual writing task. The third was used for the joint document.

Procedure

Participants arrived alone and did not meet their partner face-to-face, to ensure consistency and reduce possible effects of initial perceptions. Participants completed a pre-experiment questionnaire with demographics and prior experience with writing tools. They then viewed an instructional video on the Google Docs editing features. For CF participants, the video included instructions for the chat, comment and revision history features.

Participants then received materials for part 1 of the asynchronous task, in which they had 10 minutes to independently write an essay. Next, this document was shared via Google Docs with their partner, who was instructed to edit or provide suggestions for clarity of argument and grammar. This also lasted 10 minutes, and the document was briefly returned to the author to view the feedback. Finally, in the synchronous phase participants had 15 minutes to compose a shared document. Questionnaires were completed after each step, and full sessions lasted approximately one hour.

Measures and Analysis

Measures included behavioral logs and questionnaires.

Communication Utterances were defined as a single communication instance in one of the available channels (i.e., one complete comment or chat utterance), and copied into spreadsheets for coding. Participants in both conditions also sometimes appropriated the text editing features within the document itself for communication purposes, however. When text in the document was clearly unrelated to the content and intended to communicate with one’s partner, we included this as part of the communication log and referred to these as “in-document” utterances. Edits related to these utterances were not included in our edit logs.

Utterances were coded using a scheme based on Bales’ [1] distinction between socio-emotional and task-oriented communication. Socio-emotional and relational utterances

included those that were explicitly focused on relationships, using humor or socializing. We then divided task-focused utterances into categories identified through the coding process, with the final set described in Table 1. Coding was done independently by two coders who met after each transcript until agreement was consistently above 90%.

Type	Definition	Example
Socio-Emotional Statement	Friendship/Relationship-Building/Humor	“good working with you too!”
Review Request	Seeking approval for an action or plan	“Would that be fine?”
Agreement	Voicing support for what has been said or done	“good plan“
Disagreement	Lack of support or liking for comment or action	“I don’t think so.”
Document Evaluation	Opinion about the state of the document	“me: Good. I think it's done too”
Action Proposal	Suggesting future activities within the document	“You could delete this.”
Proposal Invitation	Asking what should be done.	“What should we do now?”
Clarification	Requesting more details or providing details.	“What does this mean?”
Task Parameters	Discussion of task details	“Are we just supposed to present the arguments but not take a side?”
Intention Statement	Indicating the speaker’s current activity	“I’m just reading this now.”
Topical Assertion	Statement about the topic of the content	“thats important cuz then the rights could fall to slater”
Softening	Using language to avoid perceptions of autonomous action	“I wrote this thing, but feel free to change it.”

Table 1. Coding categories for communication utterances.

Edits Logs of all edits were collected by manually parsing the Google Docs revision histories, following the edit division and timestamps Google provides, separated by author. Logs were coded to identify each edit as an addition or deletion, and as a “major” or “minor” edit. Major edits were defined as changes affecting one or more complete sentences, whereas minor edits affected less than one sentence. We defined a sentence as a sequence of words bounded by an initial capital letter and a final punctuation mark. The coding scheme was revised iteratively on a subset of transcripts with three coders until agreement was 100%. Subsequent transcripts were coded separately by one coder with periodic checks for consistency.

Social Attraction Modified versions of McCroskey and McCain’s [18] social attraction scales (Cronbach’s α for asynchronous phase α : .82; synchronous phase α : .89) were used to measure attraction between participants and their attraction to the task (asynchronous phase α : .74; synchronous phase α : .71). Items for both scales were aggregated

for analysis using the unweighted mean of all individual item scores. All items used 5-point Likert scales anchored by “strongly disagree” (1) and “strongly disagree” (5), such that higher scores indicate higher attraction.

Ownership and Document Quality Items were adapted from Caspi and Blau [5] with one 5-point Likert scale item measuring each of the following constructs: individual ownership, group ownership, and document quality. As above, higher scores indicate stronger ownership or quality.

RESULTS

Asynchronous Writing

This section describes results from the asynchronous task, in which participants wrote an essay independently and then made suggestions or edited their partner’s work for clarity.

Communication

RQ1a asked about the communication channels participants used to suggest edits and revisions. As there were no communication features available in the NCF condition, we expected to analyze only chat and comment logs from the CF condition. We found, however, that 25% of NCF dyads appropriated editing features for communication by making comments to their partners directly in the document text.

CF participants in this phase used only comments (and not chat) to communicate. Surprisingly, however, only 40% of CF dyads used comments. While this is more than in the NCF condition, it suggests that many participants did not perceive comments as necessary. Of those who communicated at all, the mean number of utterances for CF dyads ($M=5.43$, $SD=3.26$) did not differ from the NCF condition ($M=4.04$, $SD=2.71$) by a statistically significant margin.

We then checked for a relationship between communication and the social and task attraction variables. No differences were found, however, when we compared those from both the CF and NCF conditions who communicated at all (>0 utterances) to those in both conditions who did not.

To explore this further, we wondered if the amount of communication was related to the attraction variables. Since many communicating pairs did so only a little, it was possible that those who communicated more could differ from those who communicated less. Surprisingly, we found a moderate negative correlation ($r = -.40$, $p < .05$) between social attraction to one’s partner and one’s total number of comments for those in both CF and NCF conditions (combined) who communicated at all. This could indicate that those who felt the need to communicate more did not think highly of their partner’s work, and this affected their nascent social relationship.

To explore this further, looking only at those with at least one utterance in both CF and NCF conditions combined, we compared individuals who were above the mean ($M_{Overall}=4.32$; $SD=2.84$; $N_{Above} = 22$) number of utterances with those who were below ($N = 12$). There was no statistically significant difference between these groups for social at-

traction to one’s partner, or one’s partner’s attraction to them. There was, however, a difference in their perceived document ownership (Table 2), with below-average communicators reporting a higher individual, $F(1,31) = 16.22$, $p < .001$; and group ownership, $F(1,31) = 14.65$, $p < .01$. Frequency of communication thus had a neutral or negative relationship with both attraction to one’s partner and perceived ownership. This was surprising, so we then explored the content of participants’ communication.

	Below-Average Utterances		Above-Average Utterances	
	Mean	SD	Mean	SD
Individual Ownership**	2.68	.72	1.64	.67
Group Ownership**	3.59	.67	2.45	1.04

** $p < .01$.

Table 2. Individual and group document ownership for those with below- and above-average utterance quantities (NCF and CF combined).

We compared utterance content between the CF and NCF conditions, coded as described above. To account for variation in the number of utterances per dyad, we compared the percentage of a dyad’s total utterances in each category. As with the analyses above, those who did not communicate at all were removed from the data set for these comparisons.

Category	Overall	NCF	CF
Socio-Emotional Statement	0.79%	0%	.99%
Agreement	4.43%	7.14%	3.74%
Disagreement	1.47%	0%	1.85%
Document Evaluation	11.69%	7.14%	12.87%
Action Proposal	11.75%	37.38%	5.10%
Clarification	30.27%	10.71%	35.34%
Task Parameters	1.47%	0%	1.85%
Topical Assertion	32.92%	31.43%	33.30%
<i>Task-Oriented Total</i>	<i>99.21%</i>	<i>100%</i>	<i>99.01%</i>

Table 3. Mean fractions of utterances in coding categories for NCF and CF conditions in the asynchronous phase.

Table 3 shows the coding results. One striking outcome is that virtually all (99%) communication in both conditions was task-oriented, and not explicitly socio-emotional.

We looked at specific task-related utterances to see exactly what participants were talking about. Of all utterances, 11.75% proposed improving or changing text (“action proposal”). This meant that participants were choosing to comment rather than change the text themselves, which was particularly clear when suggested edits were minor and could have been easily made by the editor. In Case 48, for example, the editor suggests “use either ‘by anyone’ or ‘for anything,’ it would make a little more sense.” The comment provides both a simple suggestion and a rationale that implies a desire to help, thus possibly promoting positive social relations between the collaborators. Utterances in the clarification (30.27%), document opinion (11.69%), agreement (4.43%) and disagreement categories often also served

relational functions. On the whole, however, such efforts did not affect social attraction as measured here.

There was less obvious relational intent for utterances in the topic assertion (32.92%) and task parameters (1.47%) categories that were more focused on document content and parameters. Some of these may even have negatively affected attraction, as when assertions indicated different interpretations of the topic and led to disagreement.

Editing

RQ2 asked about relationships between communication and editing. Editing features were heavily used in this phase, with 96% of participants making at least one edit. There were not, however, any significant differences in editing behavior between the CF and NCF participants in editing, so we combined the conditions for these analyses.

To address RQ2a, we first examined the relationship between editing and social relationships for participants in both conditions (CF and NCF) combined. There was a moderate negative correlation between the amount of editing one did and one’s partner’s attraction to the task ($r = -.34, p < .01$). That is, the more edits one made to the document their partner wrote, the less their partner was attracted to the task. This is consistent with the result above that more communication correlated with lower social attraction, in that it suggests that people who worked in their partner’s document negatively affected task experience. Editing did not, however, appear to affect participants’ impressions of each other or the quality of the writing, as no differences were found for these variables.

	Communicators		Non-Communicators	
	Mean	SD	Mean	SD
Minor Additions*	5.26	6.41	8.30	7.64
Major Additions	.82	1.22	.94	1.12
Total Additions*	6.09	6.63	9.24	7.82
Minor Deletions	4.41	5.26	6.80	6.22
Major Deletions	.44	.96	.55	1.04
Total Deletions	4.85	5.39	7.35	6.45

* $p < .05$

Table 4. Edits by those who did and did not communicate (CF and NCF combined).

To address RQ2b, we examined the relationship between editing and communication for all (NCF and CF combined) participants with at least one utterance. On the one hand, communicating to make suggestions would indicate that these behaviors are substitutes. On the other, communicating to explain changes already made would suggest that they complement each other. We therefore compared those from both conditions with at least one utterance with those who did not communicate. As Table 4 shows, communicators made fewer additions to the document than non-communicators, $F(1,98) = 4.04, p < .05$. The same difference was not found for deletions, though is a nearly significant trend, $F(1,98) = 3.74, p < .1$.

We then looked at edits by those in both CF and NCF conditions with above- and below-average numbers of utterances, divided as described above. Supporting the idea that communication sometimes took the place of edits, several differences were found. As Table 5 shows, above-average communicators had fewer additions, $F(1,98) = 7.49, p < .01$; and deletions, $F(1,98) = 6.50, p < .05$. These differences appear to be driven by minor additions and deletions. When only major edits are compared, the differences are not statistically significant.

This suggests that those who communicated did so instead of making small changes, and particularly additions, but were more likely to make significant changes themselves; or simply not make major changes. This could be because it is easier to suggest small changes, while larger changes are harder to describe in a comment or in-document utterance.

	Above-Average Utterances		Below-Average Utterances	
	Mean	SD	Mean	SD
Minor Additions**	2.33	2.10	6.86	7.40
Major Additions	.42	.79	1.05	1.36
Total Additions*	2.75	2.26	7.91	7.51
Minor Deletions	2.25	1.71	5.59	6.15
Major Deletions	.08	.29	.64	1.14
Total Deletions*	2.33	1.67	6.22	6.22

* $p < .05$

Table 5. Asynchronous phase edits by those with above- and below average utterance frequencies (CF and NCF combined).

Our examination of transcripts yielded cases where participants both commented and edited. They often made a small edit, but explained it so their partner was aware that it occurred and why. In Case 24, for example, one partner said “you had a comma here, i took it out because it wasn’t a compound sentence.” This complicates the relationship between commenting and editing, as they can be both complements and substitutes. Importantly, though, the comment serves a similar relational purpose whether accompanied by an edit or not. Thus there was not a consistent relationship between communicating and editing in this phase.

Synchronous Writing

Communication Behavior

To address RQ1a for synchronous writing, we examined the communication logs. Unlike the asynchronous phase, all communication in this phase occurred via chat or in-document utterances. Though the comment feature was available to CF participants, none chose to use it. This suggests that chat was perceived as a way for them to coordinate effort around the document in real time, rather than comment on pieces of text. This is further reflected in the breakdown of utterances shown below.

We also saw that a larger fraction of dyads in both conditions communicated in this phase: 100% of CF dyads and 80% of NCF dyads. On average, however, CF dyads

($M=13.89$, $SD=9.57$) had over twice as many utterances as NCF dyads ($M=5.48$, $SD=4.18$), $F(1,69) = 18.56$, $p < .001$.

RQ1b asked about the relationship between communication and social relationships. We first looked at the relationship between the amount of communication and social attraction. When the CF and NCF conditions were combined and non-communicators were excluded, there was a moderate correlation between the number of utterances a participant made and their partner's social attraction to them, $r = .31$, $p < .01$. This suggests that communicating may serve a relational purpose in synchronous writing. There was not, however, a difference on these measures between those with above- and below-average numbers of utterances in this phase of the experiment for those who communicated at all in both CF and NCF conditions.

Above-average communicators combined across CF and NCF conditions, though, did report higher attraction to their partners ($M=4.03$, $SD=.38$) than did below-average communicators ($M=3.76$, $SD=.5$) by a small but statistically significant margin, $F(1,98) = 7.37$, $p < .01$. Communication also appeared to positively affect perceptions of ownership, with above-average communicators reporting a higher sense of individual ownership ($M = 3.32$, $SD = 0.70$) than did below-average communicators ($M= 2.94$, $SD = 0.94$), $F(1,98) = 4.07$, $p < .05$. There were no differences between these groups for task attraction or group ownership.

We also found that the content of communication was different here from the asynchronous phase (see Table 6). We saw more coordination and interaction, particularly in categories like "proposal invitation," "requesting review," "action proposal," and "agreement." A difference that likely had a relational effect was the greater number of explicitly socio-emotional utterances, which often made use of humor or emoticons to lighten or make the dialog friendlier. In Case 21, for example, one participant ended the session with "Alright, cool. it was nice working with you!" Sometimes these statements were used to mitigate potential relationship harm by prior actions. In Case 19, for example, one participant said "sorry for being so controlling."

Those who communicated more in this phase also tended to use more explicitly socio-emotional communication. Again combining CF and NCF participants, above-average communicators ($M =14.03\%$, $SD = 0.13$) had a larger percentage than below-average communicators ($M = 2.18\%$, $SD = 0.04$), $F(1,69) = 22.39$, $p < .001$.

In terms of task-oriented communication, there was more discussion of the task parameters in this phase, often reflecting coordination around the word limit. In Case 48, for example, one participant said "crap we have like no words left in the word count."

Again, we saw some evidence of relational intent in what was otherwise task-focused communication. Some participants used what we called "softening" statements, in which participants used language to soften potential perceptions of

their actions. In Case 42, for example, a participant made edits to the introduction, and then said "feel free to change whatever i write in the intro" to make it clear that s/he was not intending to take charge.

Further supporting this notion of relationally-oriented communication about the task, participants sometimes "invited proposals," or asked their partner for input rather than propose a specific action outright. They often began as did the participants in Case 35, with a question such as "what should we say?" Some others proposed a specific plan or took an action, but then asked for their partner to review it. While these are not necessarily relational, they do reflect some desire to involve the partner in the project. Together, these categories comprised 7.02% of utterances.

Utterance Type	Overall	Above-Average Utterances	Below-Average Utterances
Socio-Emotional Statement	7.35%	14.03%*	2.18%*
Review Request	3.22%	4.58%	2.17%
Agreement	21.87%	18.56%	24.44%
Document Evaluation	9.15%	7.08%	10.75%
Action Proposal	18.56%	15.74%	20.75%
Clarification	2.01%	1.79%	2.19%
Task Parameters	18.58%	15.77%	20.77%
Intention Statement	2.80%	3.39%	2.33%
Topical Assertion	1.62%	2.53%	0.45%
Softening	1.95%	2.50%	1.53%
Proposal Invitation	3.80%	2.25%	5.00%
Other task-related	2.76%	2.56%	1.59%
<i>Task-oriented Total</i>	<i>92.65%</i>	<i>85.97%</i>	<i>97.82%</i>
<i>*p < .001.</i>			

Table 6. Communication utterance breakdown for above- and below-average communicators in the synchronous phase.

Editing

To address RQ2a for synchronous writing, we examined the relationship between editing and participants' attraction and perceptions of the document. When both CF and NCF conditions are combined (as there were no differences between them), we found that there was a weak negative correlation between the total number of edits a participant made, and their partner's attraction to the task, $r = -.27$, $p < .01$. This suggests that, as in the asynchronous phase and in contrast to communication, editing too much can negatively affect a partner's experience of the writing process. There was, however, no similar correlation between number of edits and social attraction to or from one's partner.

Again combining CF and NCF participants, we also explored perceived ownership, comparing above-average editors to below-average. We found that those who edited more reported stronger group ownership ($M = 4.37$, $SD = .69$) than those who edited less ($M = 3.97$, $SD = .90$), $F(1,98) = 5.30$, $p < .05$. This reflects that more effort in a

document can result in a stronger sense of ownership, but the same relationship was not true for individual ownership.

Edit Type	Above-Average Utterances		Below-Average Utterances	
	Mean	SD	Mean	SD
Minor Additions*	8.74	7.55	4.68	5.37
Major Additions	3.90	2.70	3.05	2.00
Total Additions**	12.65	8.21	7.72	5.19
Minor Deletions**	9.74	8.40	3.13	3.16
Major Deletions	2.94	4.96	1.60	2.18
Total Deletions**	12.68	11.32	4.72	4.36

* $p < .05$, ** $p < .01$.

Table 7. Synchronous phase edits by those with above-and below-average utterance frequencies (CF and NCF combined).

We then looked at the relationship between the number of utterances and the number of edits. In synchronous writing, it could be the case that those who communicate a lot do so in lieu of making changes, and make suggestions instead. On the other hand, it could be that more communication leads to more discussion about possible edits, and therefore more edits. Our analyses support the latter explanation, showing a moderate positive correlation between the number of utterances and the number of edits for participants in both CF and NCF conditions combined, $r = .36$, $p < .001$.

Moreover, as Table 7 shows, those who were above-average communicators in the CF and NCF conditions combined had more additions and deletions than did the below average communicators, both by statistically significant margins (Additions: $F(1,98) = 8.03$, $p < .01$; Deletions: $F(1,98) = 4.47$, $p < .05$).

Edit Type	Above Average Edits		Below Average Edits	
	Mean	SD	Mean	SD
Minor Additions	39%	15%	38%	29%
Major Additions**	15%	10%	32%	26%
Total Additions**	54%	15%	70%	20%
Minor Deletions**	38%	12%	20%	17%
Major Deletions	9%	10%	10%	14%
Total Deletions**	46%	15%	30%	20%

** $p < .01$.

Table 8. Synchronous phase edits by those with above-and below-average edit frequencies (CF and NCF combined).

We next wondered whether those who edited more than average did so differently than those who edited less. We therefore compared above-average editors to below-average, again combining CF and NCF participants. For these comparisons we look at percentages of edit types, rather than raw counts, as counts would clearly be higher for those with more overall edits. We first found that below-average editors had a higher percentage of additions than did the above-average editors (Table 8), $F(1,98) = 16.77$, $p < .001$. Breaking these down further, we found that these differences were driven in particular by major addi-

tions ($F[1,98] = 13.38$, $p < .001$) and by minor deletions ($F[1,98] = 27.42$, $p < .001$). This suggests that above-average editors made a lot of small deletions, while below-average editors tended to make major additions.

DISCUSSION

We began with questions about socio-emotional and relational communication in collaborative writing. While tools for writing have traditionally been designed primarily for providing awareness of what others are doing or have done in a document [28], our results provide more evidence that people also engage in communication for relational purposes when writing together.

Implications for Theory

From a theoretical standpoint, we believe our results argue for more consideration of relational and group maintenance behavior in discussions of awareness.

First, we saw that communication around documents was important, whether or not participants were provided with chat/comment features. More CF participants communicated in both asynchronous and synchronous writing tasks, but many NCF participants appropriated the text editing features to use them for communication. Despite the importance of communication, however, its relational effects were mixed. In the asynchronous phase, there was a negative relationship between the number of utterances and both attraction to one's partner and ownership of the document. In the synchronous condition, on the other hand, more communication was generally associated with higher levels of attraction and ownership.

This suggests first that communication is particularly important from a relational standpoint when participants are working directly together, and possibly when they have an existing relationship. The asynchronous result has several possible explanations. Commenting or in-document utterances, where present, may have annoyed people or been seen as intrusions in a document that felt individual. Where no comments were used, participants may not have felt anything was missing, so there was no effect.

It is also possible that our findings in the asynchronous phase are partly an artifact of the study design, in which pairs of strangers collaborated. The need to comment or the presence of comments from others could have been more annoying than with a known partner. This result could have been different with pairs known to each other or if the pairs had been given time to interact (as in [29], in which chat sessions helped build trust) prior to this part of the task.

People's choices of communication channels also reflected the nature of their task, in addition to what was available to them. NCF participants could only communicate by appropriating text editing. CF participants communicated almost exclusively via the commenting feature in the asynchronous phase and exclusively via chat in the synchronous phase. Chat was arguably more conducive to real time coordination, but it also seemed to be more conducive to explicitly

socio-emotional utterances. These occurred more frequently in the CF than NCF condition.

Second, we saw variations in editing behavior that were related to the number of utterances. Those who communicated less tended to edit more in both conditions, particularly for minor additions and deletions. Major changes did not differ between above- and below-average communicators, likely because these were less likely to be described easily in a comment or chat utterance.

Third, editing and communication were related to social factors, but this relationship was not always straightforward. In the asynchronous phase, both the amount of editing and commenting one did were negatively correlated with attraction measures. This may mean that, in asynchronous writing, there was less of a perceived relationship with one's partner. The need to comment on their work or have one's work heavily edited may have been seen as a burden or intrusion. Ownership results support this in that group ownership in the asynchronous phase was below the scale midpoint and lower for below-average communicators.

In the synchronous phase, there was a positive relationship between the number of utterances and the social attraction of one's partner. There was also a relationship between communication and ownership; though in the synchronous phase this was only true for individual ownership. Group ownership, however, was related positively to the amount of editing one did in the document; while editing was also related negatively to one's partner's attraction to the task. Thus, even in a synchronous document there may have been some perceptions of invasiveness or burdensome edits.

Implications for Design

Our results also have implications for the design of future systems to support collaborative writing. First, our results show evidence that communication around documents is valued and can have both positive and negative socio-emotional and relational effects. This suggests that designers should think not just about how to make people aware of changes and what others have done in a document, but also give them a chance to show evidence of positive intent. This could be accomplished via language, as we saw our participants do, as well as via targeted features.

Consistent with design ideas presented by Birnholtz and Ibara [4], these results further point to the utility of features to support the presentation of revisions as possibilities for the future state of the shared document, rather than as definitive changes. This could be very accomplished, for example, by marking these as “<Name> suggests deleting <text>” instead of showing the deleted text. The system could also provide an easy way to explain the reason for an edit that could be incorporated in this. Consistent with Goffman [12] and theories of accounting [6], these features would enable people to present their intent as positive.

Second, we saw that participants who had larger numbers of utterances tended to make more minor edits, while the

number of major edits did not differ as much by communication frequency. The first implication of this is that features for explaining and presenting edits can focus on minor edits – that can more easily be represented and highlighted. The second implication, though, is that designers may wish to consider tools to support easier description and explanation of more major edits and restructuring. One possible way to do this would be via a “fork” in the document structure (as with code repositories, for example) or a “layers”-type interface as used in Adobe Photoshop, for example.

We also saw our participants communicate almost exclusively via comment in the asynchronous phase and via chat in the synchronous phase. Even in chat, though, we saw them discuss specific elements of the document; even though only comments can be clearly tied to specific text fragments. This suggests that it may be useful for chat utterances to be tied to specific text elements as well.

Limitations and Future Work

As with any experiment, there are tradeoffs between external validity and experimental control. To avoid possible external effects of prior relationships, our participants were paired with strangers. While this enabled us to test relational effects with minimal extrinsic bias, it does mean that participants may not have engaged in as much relational behavior as they might with a real-world collaborator. We believe, however, that seeing any evidence of relational effects in this study suggests that they could be even more pronounced where there was a higher likelihood of working together again. This urges interpretation with caution, however, and is a key topic for future work.

Two limitations urge further caution in interpreting our results. First is our two-phase, open-ended task. The advantage of this task is that it allowed us to simulate parts of a real writing process, in which participants were wrote real documents that would be evaluated. A limitation, however, is that one was editing the work of a stranger. The two phases also conflate two related independent variables: synchrony/asynchrony, and individual/group documents. While informative, this design does not allow us to tease apart possible differences that might occur if people were to synchronously edit an individual's document, or asynchronously work on a document that is more clearly shared. The second limitation concerns our statistical analyses at the individual level of analysis, which could be affected by the non-independence of dyadic observations.

An unexpected limitation is that participants did not use the revision history features at all. This likely reflects that they were working on a short document with only one other person, so could likely see changes and knew that, if they didn't edit it, their partner probably had. Future work could usefully include a design with a longer document and/or more participants in a scenario where the revision history information could be more helpful.

ACKNOWLEDGMENTS

We thank Google for its partial support and acknowledge valuable assistance from Rachel Boochever, Madeline Bourgeois, Ryan Jones, Charles Lai, Jenni Lieberman, Kate Pascucci, Lindsay Reynolds, Behszod Sirjani and Emma Zuroski. We also thank Judy Olson for her feedback.

REFERENCES

1. Bales, R. F. *Interaction process analysis: a method for the study of small groups*. Addison-Wesley, Cambridge, MA, 1950.
2. Beck, E. E. and Bellotti, V. Informed opportunism as strategy: supporting coordination in distributed collaborative writing. In *Proc. ECSCW* (1993), 233-248.
3. Bietz, M. J. Effects of communication media on the interpretation of critical feedback. In *Proc. ACM CSCW* (2008), 467-476.
4. Birnholtz, J. and Ibara, S. Tracking Changes in Collaborative Writing: Edits, Visibility and Group Maintenance. In *Proc. ACM CSCW* (2012), 809-818.
5. Caspi, A. and Blau, I. Collaboration and psychological ownership: how does the tension between the two influence perceived learning? *Social Psychology of Education*, 14, 2 (2011), 283-298.
6. Cody, M. J. and McLaughlin, M. L. Interpersonal accounting. In H. Giles and P. Robinson, ed. *Handbook of language and social psychology*. Wiley and Sons, London, 1990.
7. De Croon, E. M., Sluiter, J. K., Kuijer, P. P. F. M., and Frings-Dresen, M. H. W. The Effect of Office Concepts on Worker Health and Performance: A Systematic Review of the Literature. *Ergonomics*, 48(2005), 119-134.
8. Dourish, P. and Bellotti, V. Awareness and Coordination in Shared Workspaces. In *Proc. ACM CSCW* (1992), 107-114.
9. Ellis, C., Gibbs, S. and Rein, G. Groupware: some issues and experiences. *CACM*, 34, 1 (1991), 39-58.
10. Erkens, G., Jaspers, J. and Prangsa Coordination processes in computer supported collaborative writing. *Computers in Human Behavior*, 21(2011), 463-486.
11. Goffman, E. *Behavior in public places*. The Free Press, New York, 1963.
12. Goffman, E. *The presentation of self in everyday life*. Anchor Books, New York, 1959.
13. Gutwin, C. and Greenberg, S. A descriptive framework of workspace awareness for real-time groupware. *CSCW*, 11(2002), 411-446.
14. Hackman, R. The design of work teams. In J. Lorsch, ed. *Handbook of organizational behavior*. Prentice-Hall, Englewood Cliffs, NJ, 1987.
15. Ignat, C. L. and Norrie, M. C. Multi-level editing of hierarchical documents. *CSCW*, 17, 5-6 (2008), 423-468.
16. Kim, H. C. and Eklundh, K. S. Reviewing practices in collaborative writing. *CSCW*, 10(2001), 247-259.
17. Leland, M. D., Fish, R. S. and Kraut, R. E. Quilt: A collaborative tool for cooperative writing. In *Proc. ACM CSCW* (1988), 206-215.
18. McCroskey, J. C. and McCain, T. A. The measurement of interpersonal attraction. *Communication Monographs*, 41(1974), 261-266.
19. McGrath, J. E. *Groups: Interaction and Performance*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
20. Neuwirth, C. M., Chandhok, R., Charney, D., Wojahn, P. and Kim, L. Distributed Collaborative Writing: A comparison of spoken and written modalities for reviewing and revising documents. In *Proc. ACM CHI* (1994), 51-57.
21. Neuwirth, C. M., Chandhok, R., Kaufer, D. S., Erion, P., Morris, J. and Miller, D. Flexible diff-ing in a collaborative writing system. In *Proc. ACM CSCW* (1992), 147-154.
22. Neuwirth, C. M., Kaufer, D. S., Chandhok, R. and Morris, J. H. Computer Support for Distributed Collaborative Writing: A Coordination Science Perspective. In G. M. Olson and T. W. Malone, ed. *Coordination Theory and Collaboration Technology*. Lawrence Erlbaum Assoc, Mahwah, NJ, 2000.
23. Noel, S. and Robert, J.-M. Empirical study on collaborative writing: what do co-authors do, use and like? *CSCW*, 13(2004), 63-89.
24. Olson, J. S., Olson, G. M., Storrosten, M. and Carter, M. Groupwork close up: A comparison of the group design process with and without a simple group editor. *ACM Trans. on Info. Systems*, 11(1993), 321-348.
25. Posner, I. and Baecker, R. How People Write Together. In *Proc. HICSS* (1993), 127-138.
26. Posner, I., Mitchell, A. and Baecker, R. Learning to write together using groupware. In *Proc. ACM CHI* (1995), 288-295.
27. Scott, M. B. and Lyman, S. M. Accounts. *American Sociological Review*, 33(1968), 46-62.
28. Tam, J. and Greenberg, S. A framework for asynchronous change awareness in collaborative documents and workspaces. *International Journal of Human-Computer Studies*, 64(2006), 583-598.
29. Zheng, J., Veinott, E., Bos, N., Olson, J. S. and Olson, G. M. Trust without Touch: Jumpstarting long-distance trust with initial social activities. In *Proc. ACM CHI* (2002), 141-146.