The Multiple Views of Inter-organizational Authoring

David W. McDonald
The Information School
University of Washington
Suite 370, Mary Gates Hall
Seattle, WA 98195
dwmc@u.washington.edu

Chunhua Weng
Biomedical and Health Informatics
University of Washington
1959 NE Pacific St. Box 357240
Seattle, WA 98195
cweng@u.washington.edu

John H. Gennari
Biomedical and Health Informatics
University of Washington
1959 NE Pacific St. Box 357240
Seattle, WA 98195
gennari@u.washington.edu

ABSTRACT
Collaborative authoring is a common workplace task. Yet, despite improvements in word processors, communication software, and file sharing, many problems continue to plague co-authors. We conducted a qualitative study in a setting where participants are loosely connected, physically separated, and work together over a period of 4-9 months to author a complex technical document—a clinical trial protocol. Our study differs from most prior work in that the collaboration is longer-lived, and that the collaborators do not share equivalent status, background, nor domains of expertise. Our data demonstrates that the participants do **not** share the same view or representation of the authoring process, even though it has a long organizational history. Nonetheless, the participants can still coordinate their activity while maintaining only partially consistent representations of what they are doing. We contend that partial consistency in the participants' concept of the collaborative process is a feature for their asynchronous collaboration at a distance. Based on our findings we suggest a number of improvements for both tools and tool usage that have direct impact on support for collaborative authoring.

Categories and Subject Descriptors
H.4.1. Office Automation: Groupware; H.5.3 Group and Organization Interfaces: Computer-supported Cooperative Work.

Keywords
Collaborative authoring, field study, clinical trials, medical informatics.

1. INTRODUCTION
Collaborative authoring has historically been fertile ground for research and development in Computer-Supported Cooperative Work (CSCW). Studies of groups collaboratively writing have varied from grade school children in a classroom [9] to MBAs in a laboratory [12]. Some studies are more self reflective of the writing with new tools [7] while other studies have been facilitated with broader system deployment [11].

Our study builds on this tradition by considering a specific type of collaborative authoring that is done at a distance. This research focuses on clinical trial protocol authors. A *clinical trial protocol* is a highly-structured, self-contained document that describes all of the procedures required to carry out clinical medical research, including information about patient enrollment, data collection from participants, and statistical information and endpoints for analysis.

CSCW'04, November 6–10, 2004, Chicago, Illinois, USA.

This type of collaborative authoring is different from that of many prior studies. In particular the complexity of the developing document often requires more participants with a wider range of expertise than most. Secondly, the development of a clinical trial protocol document often takes more than six months. As such the collaboration is more long-lived than that in many studies of collaborative authoring.

We conducted a qualitative study to understand the collaborative process of a clinical trial protocol authoring. We studied the activity through artifact collection, observations, and interviews. During several interviews participants were asked to diagram their view of the authoring process. These diagrams were used to frame follow-up questions and understand the way this activity is coordinated.

We focus here on elaborating the varied representations of the authoring process. The processes become a way to reflect on the asynchronous coordination necessary to complete a collaboration which spans multiple months. In general, we find that participants do not converge on similar representations of the process, but that they can still coordinate their activity to generate a final product. The varied motivation of the participants and their varied levels of intensity in the collaborative effort are effectively supported through differing personal representations of the process.

Our motivation for this study is to address the needs of clinical trial protocol authors through the design of tools to support their collaborative activity [21]. The organization in the study has been looking to improve the efficiency of the authoring process. However, based on our results they do not see the need for wholesale process re-engineering. Rather, and with our encouragement, they see the need to develop more flexible tools and systems that could result in smaller work practice changes to improve communication, coordination, and document management practice.

This paper details our findings about the various conceptions of the collaborative authoring process, highlighting similarities and differences. This is used to consider how different conceptions of the authoring process can facilitate coordination. Lastly, we consider possible improvements in the collaborative process, but not as traditional process reengineering. We frame the improvements with respect to existing tools and the development of a new tool to support the processes in the organization.

2. AUTHORING IN THE FIELD
Collaborative authoring studies have been conducted with a range of methods. A number of studies have focused on the writing activity as supported by a specific collaborative

authoring tool. Examples of this type is the studies by Mitchell et al. [9], Neuwirth et al. [11], and Olson et al. [12]. Many of these studies rely on a specific editor (Aspect, ShrEdit), a specific task (write a magazine article, create a design document), a specific set of participants (youth, MBAs), and the collaboration evolves over a relatively short period of time (a few hours, to 12 weeks).

Other studies focus on writing activity unassisted by a specific tool or collaborative editor. In these studies the goal is to understand how collaborative authoring is conducted while relying on general communication and coordination technologies (e.g., phone, email, word processor, ftp, etc.) .

Some early survey research pointed to problems in defining collaborative authoring. In a survey of authors Beck [1] found more than two-thirds of collaborative authoring efforts rely on a single primary author. Dillon [3] focused on the collaborative writing processes of two technical report teams. Dillon confirmed Beck's finding that more than two-thirds of collaborative authoring efforts rely on a single primary author. He found that both teams relied on a primary author to generate an initial draft, which results in "a single authoring seeking group feedback" type process.

Elaborating the collaborative authoring process further, Posner and Backer [14] conducted 10 interviews to investigate the collaborative writing in 22 different projects. Through their interviews they identified a range of roles, writing strategies, writing-related activities, and document-control methods. Some studies do not present a rich, elaborated view of the authoring process, but instead create an almost prescriptive framework for engaging collaborative authoring. One set of stages includes analytic, synthetic, problem solving and evaluation [16]. Another set of stages includes planning, translating, and reviewing [5]. While a third study identified evaluating the writing situation, conducting research, writing the document, reviewing and revising as key stages [13].

Building on Beck's initial survey Beck & Bellotti [2] noted that there are few longitudinal studies of collaborative authoring. Beck & Bellotti conducted *in vivo* case studies of three collaborative authoring projects that lasted between three weeks and one year. The three groups relied heavily on phone, diskette, fax, and 'electronic file exchange' to communicate and exchange the shared document.

Beck & Bellotti found great variety in writing practices. The design recommendations were not explicit to the authoring process. Instead, design recommendations focused on the coordination of the editing activities (e.g., edit history, formatting, commenting, collaborative tools, etc) and the coordination of responsibilities among the co-authors (e.g., planning and status, progress information, access control, etc).

Our study is in the tradition of that by Beck & Bellotti [2]. Our study differs from theirs in that we focus on co-authors conceptions of the writing process. Additional problems of collaboration are framed relative to the processes that our participants define. As well, all of our users rely heavily on email, which is only partially true in Beck & Bellotti's study.

Collaborative authoring is common in healthcare settings. However, it has not been widely studied and is not well supported. A number of recent studies examine the development of clinical trial protocols. The focus has been on improving the processes through automation. These efforts can fall into the following three categories: (a) use of a domain ontology and computer-based decision support tools to guide clinical protocol authoring, [10] (b) use of mark-up languages or models to transform existing free text clinical protocols into a computer-interpretable format and provide critiques to the protocol document content, [15] and (c) use of automated processing techniques to extract reusable components of previous clinical trial protocols [17] and provide templates to facilitate knowledge reuse in clinical trial design [4]. An underlying assumption for most of this work is that a single person writes the protocol and as such sophisticated tools are necessary to overcome the gaps in a person's capacity and knowledge.

Our study departs from that assumption in the current medical informatics literature by considering the writing process to be a collaboration among a diverse group of experts. As such, our study is in the tradition of CSCW studies of collaborative authoring. Our work builds on prior studies, such as that by Beck & Bellotti [2] by examining the authoring processes that evolve over a minimum of four months. And, while each clinical trial protocol is unique in some way, several of our study participants collaborate on the development of many protocols each year.

3. THE FIELD SITE AND DATA

The field site is not a single organization in a single location. The collaborative activity of authoring clinical trials is coordinated through umbrella organizations. An umbrella organization is the hub of activity, but in general, participation crosses organizational boundaries. The various participants and the organizations from which they come have differing goals with regard to a clinical trial. These differing goals lead them to focus on different aspects of the protocol document and attend to the collaborative activity of authoring with different levels of intensity.

3.1 Southwest Oncology Group

The organization that frames our study of collaborative authoring is the Southwest Oncology Group (SWOG). SWOG is a consortium of institutions and investigators chartered and funded by the National Cancer Institute (NCI) for improving the survival of cancer patients through clinical research. SWOG was started by the NCI during the 1950's and is headquartered in San Antonio, Texas (in the southwest part of the United States). SWOG is one of twelve consortia that develop, conduct, and analyze complex cancer research trials.

One of SWOG's primary tasks is to help develop clinical trial protocols. A clinical trial protocol is an integration and representation of multidisciplinary expertise. This expertise comes from physicians, drug companies, biostatisticians, and protocol coordinators. In general, a physician (the principal investigator) will come up with an idea for a clinical trial and write between one paragraph and one page on the idea. However, this individual does not have the required statistics training to understand issues of protocol design and statistical significance. Thus, SWOG coordinates and connects principal investigators with appropriate staff biostatisticians. In addition, there is a substantial body of institutional knowledge about safety issues and document standards that are critical for protocol approval. This knowledge is provided by SWOG through biostatisticians and a protocol coordinator. Over a period of 4-9 months a team minimally including a protocol coordinator, a principal investigator, and a biostatistician will produce pro-

tocol document of 60-120+ pages. SWOG collaborators write about 40 clinical trial protocols each year.[1]

Clinical trials are complex and expensive and the protocol is the starting point. SWOG has a vested interest in making certain that a protocol will meet strict standards for both patient safety and scientific and statistical significance. There are strong economic incentives to make certain that SWOG does not author protocols that do not meet these requirements. A protocol will go through multiple rounds of peer review. SWOG arranges peer-reviews as a means of improving a protocol and as a way of identifying problems, inconsistencies, and oversights in a trial. These reviewers include experts from outside organizations and from a range of disciplines. The first level review is a Protocol Review Committee (PRC), consisting of other SWOG employees and researchers. Comments from this level of review are often distributed to experts from other organizations. At the national level, the Cancer Therapy Evaluation Program (CTEP), within the NCI, reviews the protocol. As well, a local Institutional Review Board (IRB) will review every protocol at each institution that participates in the trial. SWOG protocol development processes make it easier to design and carry out multi-site protocols. SWOG designs trials to address the concerns of as many local IRBs as possible.

When developing and managing a protocol, the problems that SWOG must overcome are not just the management of diverse expertise. The effort is geographically dispersed—rarely, if ever, are all collaborators on a protocol in the same physical locale. SWOG is a geographically distributed organization, with protocol coordinators at the headquarters in Texas, biostatisticians at the University of Washington in Seattle, and principal investigators scattered at affiliated institutions across the western United States. Any given protocol coordinator and statistician will participate in authoring many protocols each year, but a principal investigator will work on far fewer.

The principal investigator (PI) is the source of the idea for the clinical trial, and is considered the ultimate authority on its design. The inspiration for a trial might come from a symposium, a paper, or a relationship with a drug company researcher. A PI is generally a practicing physician at an academic medical center. A single, successful clinical trial of the type we studied can make the entire career of one physician. It is rare for a PI to lead more than one or two of these multi-site trials in an entire career.

While the PI might be the source of the idea, the PI is not the complete focus of all activity. In addition to the PI, the collaboration to develop a clinical trial will often include specialists who minimally provide comments at regular stages. For example, if a drug is involved, often a specialist from the drug manufacturer will participate. Pathobiologists, nursing specialists, data coordinators, and senior SWOG managers will all provide input or modify the document at different stages.

A clinical trial can be a source of great pride and success, but also great heartbreak. Many clinical trial protocols traverse a significant number of hurdles but are never conducted because of some tragic circumstance. For example, a drug company might discover that it is too costly to produce an experimental drug in the quantities needed for a large, statistically significant trial.

We use SWOG as a loose boundary for the study, but like all studies of medical institutions, practicing physicians play critical roles in the process. Because physicians are pulled in so many directions, having them as participants can be difficult. The effort in these cases is to get the best data possible from a small number of participants in the short times that are available for interacting, observing, and interviewing.

3.2 Methods and Data Collection

The research was conducted using standard qualitative methods. The primary data includes informal and semi-structured interviews, artifact collection (e.g., protocols, email exchanges, meeting minutes, protocol review comments), and observation. We also attended PRC review meetings from April 2002 through June 2002 and from November 2002 to April 2003. Overall we observed 30 protocol reviews, where some protocols were reviewed multiple times.

Participants were identified through snowball sampling. Interviews with a SWOG statistician resulted in an interview with a PI and a SWOG protocol coordinator who referred us to others. We conducted nine in-person semi-structured interviews with clinical trial protocol authors; two biostatisticians, three protocol coordinators, three principal investigators and one protocol coordinator manager. Each interview lasted about an hour and included some common questions and two activities, discussed below. Additionally, there have been many informal interviews and interactions with senior SWOG managers during the past two years.

During the semi-structured interviews we inquired about many aspects of clinical trial protocol authoring (e.g., previous protocol authoring experience, roles, how efforts are coordinated, how document revisions are handled, communications, etc). Since email is a primary communication and coordination mechanism, we were interested in how it was being used by groups of protocol authors.

We asked each interviewee if we could be included in the communication circles of a developing protocol. It was the responsibility of the participant to pick which protocol (if any) and provide an initial introduction. We followed up with others in the collaborative group to get their approval. We tracked two protocols through email communication and have a third currently being tracked. The protocols were all at different stages, but most were early in the process. The emails included a number of attachments that are specific to protocol development, including protocol drafts, review comments, policy and standards, and protocol authoring guidelines.

Our interviews included one or two activities depending on how much time we had. Loosely we called these activities a *document review* and a *process query*. We will not elaborate the document review method, but it was useful in getting participants to think about the protocol authoring process.

3.2.1 Process Query

Different participants see a collaboration process in different ways. Suchman [18] has pointed out that to an outsider, work always seems less complex than it actually is. Participants who collaborate at a distance have a similar problem. Each participant only sees a portion of what is actually going on.

As a second activity, we asked the interviewee to draw a graph or picture of the protocol authoring process. Participants natu-

1. SWOG also helps conduct the clinical trial in association with member institutions. At any one time, SWOG manages about 120 clinical trials in various stages.

rally detail aspects of the collaboration that are most salient to them. The participants' specific diagram was then used to frame follow-up questions that probe portions of the process which the participant highlighted or omitted. For example, in portions of a diagram with high detail, we asked the participant to provide an example where the process worked particularly well or where it worked poorly.

Some of the most informative data came from participants' efforts to describe the protocol authoring process. The following analysis focuses on the process of clinical trial protocol authoring. The representations of the process and how participants elaborate the process frame problems they have communicating with each other, problems with current collaboration tools, and problems coordinating activity.

4. THE VIEWS OF A COLLABORATIVE PROCESS

Like many complex organizations, process is important to SWOG. Process is both the 'how' and the 'what' of the work for SWOG. It is the 'what' because clinical trials are themselves very complex processes, with many steps and stages. Writing a clinical trial protocol is partly the task of specifying a complex process in excruciating detail. But process is also the 'how' of SWOG because, as a complex organization, their own process is how they get their work done.

Through this analysis we will describe the collaborative authoring process from three points of view. First, we will consider the process as represented by the organizationally sanctioned training materials. Then we will consider the process from the perspective of a principal investigator (PI). The PI view is interesting because a PI might write only one or two clinical trials in a career. Last, we will look at the process from the point of view of SWOG protocol coordinators.

4.1 The Official SWOG Process

SWOG maintains a set of policy documents that describe in a rational way how the organization works. One of the more complex documents that SWOG maintains is titled *Protocol Guidelines*. This is a large document that describes how a clinical trial protocol is structured, and how it is processed. SWOG, like many organizations, must deal with employee turnover. The Protocol Guidelines facilitate training of both new SWOG employees and new PIs. Once or twice a year, a small number of physician researchers are selected for SWOG training. These researchers visit the SWOG statistical center, and participate in a PRC meeting that reviews their own protocol ideas. This training opportunity is usually the only time that a prospective PI ever sees the process in action.

The Protocol Guidelines document includes a complex, rationalized, view of the protocol process (see Figure 1). The process is represented in a classic Program Evaluation and Review Technique (PERT) chart. The specific text in each box of the process is not important—rather, we highlight some of the structural features of the chart.

On the left side of the diagram, the process begins with a small collection of five boxes that lead into the more complex middle of the diagram. These steps in the process are concerned with developing and rating a *capsule summary*—a brief 1-2 page summary of the entire protocol (usually authored by the PI). Disease specific committees prioritize these summaries, and the highly-rated proposals are moved through the remaining phases as quickly as possible.

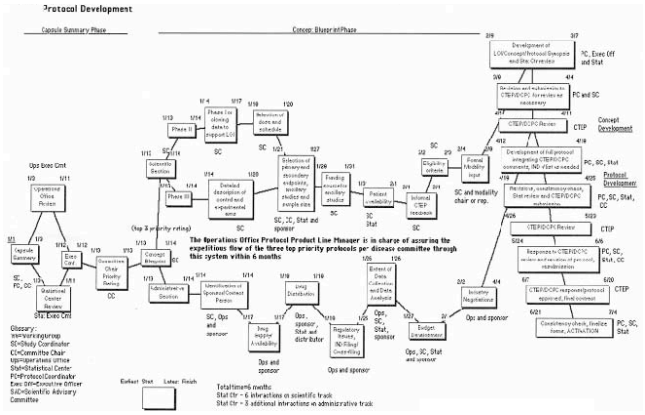

Figure 1. Protocol Process from SWOG Protocol Guidelines

The middle of Figure 1 shows a split into three parallel tracks of tasks that become two parallel tracks. This middle phase results in a complete draft of a clinical trial protocol. The two parallel tracks correspond to administrative and intellectual development tasks of a protocol document. Finally, the right hand side of the figure shows a vertically cascading set of tasks. These tasks include submission to and review of the protocol by the Cancer Therapy Evaluation Program (CTEP). This is the final step before the activation of the clinical trial.

Each task in the PERT chart is labeled with starting and ending dates and responsible personnel. Based on the frequency of labeling, the key personnel in the overall process are the PI and SWOG's statisticians. A close second is a SWOG protocol coordinator. The dates on the chart have been carefully scheduled so that from beginning to end a clinical trial protocol could be produced in six months.

Figure 1 conveys a relatively complex, but still, very 'rational' process of protocol development. For example, the chart does not include any protocol review committee (PRC) meetings. As well, some tasks require the completion of specific sections of the protocol. These tasks occur in advance of PRC meetings. As a result of a PRC meeting it is fairly common for the reviewers to require significant changes to a protocol. Thus, the real process has some cycles which are not represented.

Still, as an organizational tool, Figure 1 serves an important need to communicate the process to new employees and to prospective SWOG PIs. However, in the following we see that PIs actually come to understand the process quite differently.

4.2 The PI View of the Authoring Process

Each of the PIs we interviewed had authored at least one successful protocol. One of our more senior PIs was getting ready to author his second. While each had different approaches to managing the complexity of the collaboration, they had some strikingly similar views of the protocol authoring process. Figure 2 and 3 are images of the process diagrams that two of our PIs produced.

All three PIs began viewed the process as starting before the writing of a capsule summary. In Figure 2 the node 'idea' starts the process. In that Figure, the idea is influenced by the stats group and by a disease committee. The PI felt that this early verbal approval was important to getting a capsule summary started in the right direction. The PI drawing Figure 3 specifically highlighted three sources for ideas, which lead into an interaction with a disease committee. Those three sources include 'new technology' (e.g., new drugs), 'previous protocols'

that were developed by SWOG, and 'protocols' that were non-SWOG. The PIs were much more concerned with the idea generation phase of protocol authoring.

All of the PIs highlighted the importance of feedback and cycles in the protocol writing process. In Figure 2 above the cycle consists of the activities in the circle in the middle right of the process diagram. Figure 3 shows a pair of circular arrows, one labeled 'revisions.'

Another similarity among all of the PIs process diagrams was the importance and centrality of the SWOG biostatisticians. All of the PIs included 'stats' early in the process. This is probably because of the complexity of structuring multi-arm[2] clinical trials to ensure statistical significance. Most PIs do not have the specialized stats background to completely frame the experimental structure.

It is also notable that none of the PIs include a Protocol Coordinator (PC) as part of their process. There are references to using services from SWOG 'ops' (operations) to get templates or other services, but nothing specific about the PC. This is striking because of the very central role that a PC plays in the actual writing.

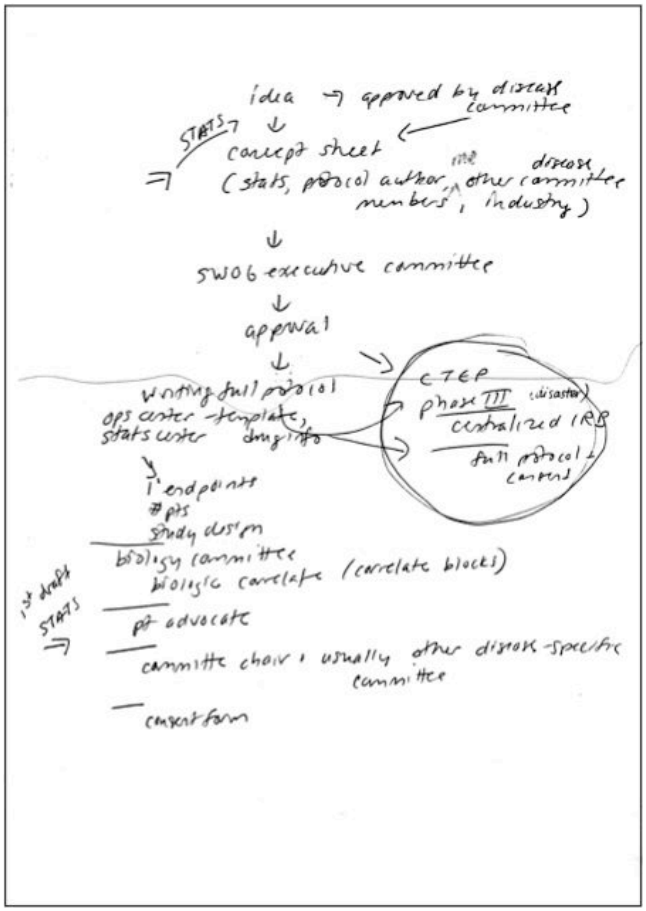

Figure 2. Process as viewed by principal investigator 1.

We were included in the email communications on two protocols under development. From these emails, we see the important coordinating role played by the PC. In some ways, the PC actually drives the process, reminding the PI to submit crucial sections of the document, organizing the peer-review sessions with the PRC, clarifying the feedback and sending to the PI, getting the PI to respond to the feedback, as well as editing and managing the drafts of the protocol.

It may be unfortunate that PIs overlook the PC, but it is interesting because of the role the PC plays in the process. The PC is the main organizer and driver of the multi-track 'middle' of Figure 1. All of the PIs' process diagrams overlook the complexity in that stage. One PI stated that after submitting a capsule summary, he would wait several months, and magically, a complete protocol would show up in email. That PI later elaborated the process, but it is clear that the official SWOG view is not accurate with respect to the level of PI involvement. The PCs take care of much behind the scenes.

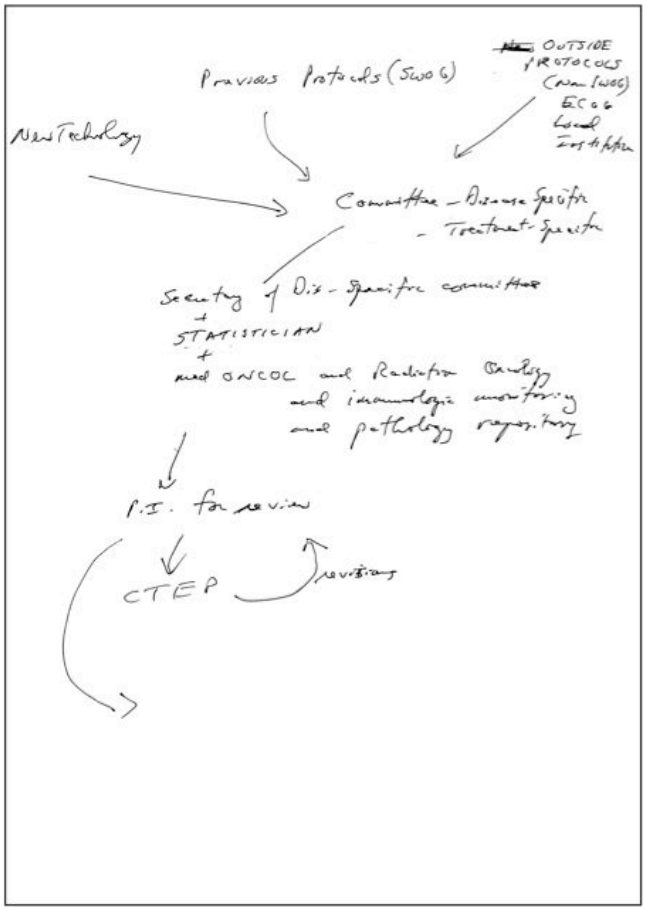

Figure 3. Process as viewed by principal investigator 2.

4.3 The PC View of the Authoring Process

SWOG maintains a number of permanent staff to assist with the writing of clinical trial protocols. One key staff member is the Protocol Coordinator. SWOG trains each PC using the same materials as provided to prospective PIs. While PIs are provided an introduction to the process, a new PC is carefully mentored by an existing PC. This is feasible because all SWOG PCs are collocated at SWOG headquarters.

PCs are responsible for shepherding protocols through the development process. PCs work on several protocols at the same time. A PC contributes to the development by writing

2. Multi-arm is the biostatisticians term for an experimental design with multiple experimental conditions or cohorts.

sections, providing boilerplates, developing supporting forms (e.g., informed consent forms), and consistency checking among sections. The PC also sets the pace of collaboration and functions to coordinate much of the collaborative activity. PCs schedule peer reviews with the Protocol Review Committee (PRC) and the Clinical Trial Evaluation Program (CTEP). PCs are responsible for collating feedback from the reviews and for getting the PI to respond to the feedback.

We interviewed 3 PCs and the SWOG PC manager. At the time of our interview one PC was relatively new and had only been on the job about 1 month. The others had one or more years of experience and participated in many protocols. The view of the development process among the PCs was much more varied. In Figure 4 we show the view of the newer PC, and in Figure 5 we show the view of a very experienced PC. This is the extreme case, but the point here is to reflect the variation in the views.

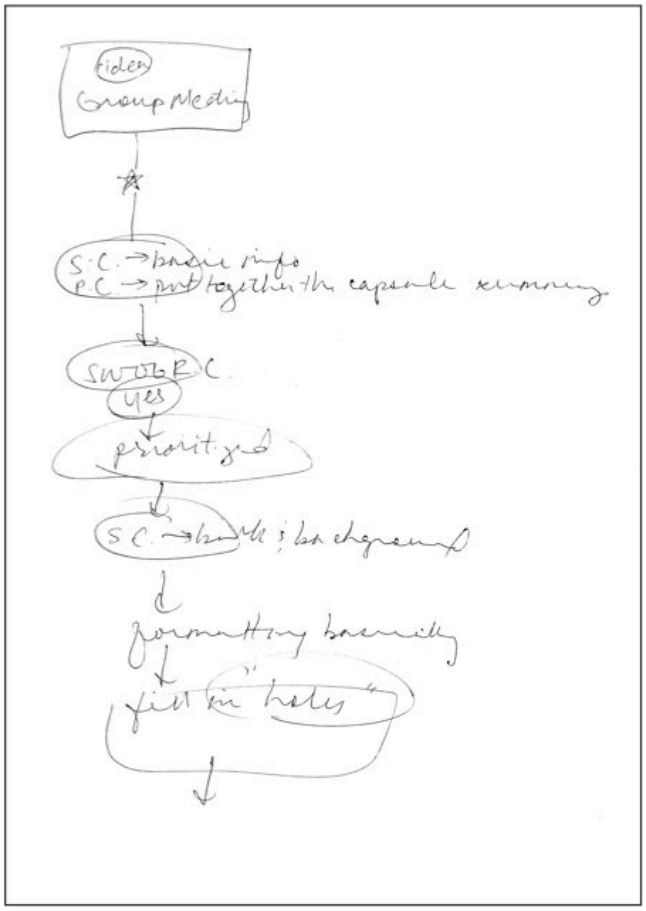

Figure 4. Process as viewed by a new coordinator.

Both views start with having a PI propose an idea.[3] The PCs generally consider themselves involved early in protocol development. The novice PC included involvement in the development of the capsule summary (e.g., process in Figure 4), whereas the experienced PC started involvement after the priority ratings of the summary (see Figure 5). The novice PC process diagram shows a linear process, whereas the experienced PC shows feedback cycles, like Figure 5. Interestingly, both process diagrams show little involvement by the SWOG statistical group. This is different from the official view that heavily involves the stats group.

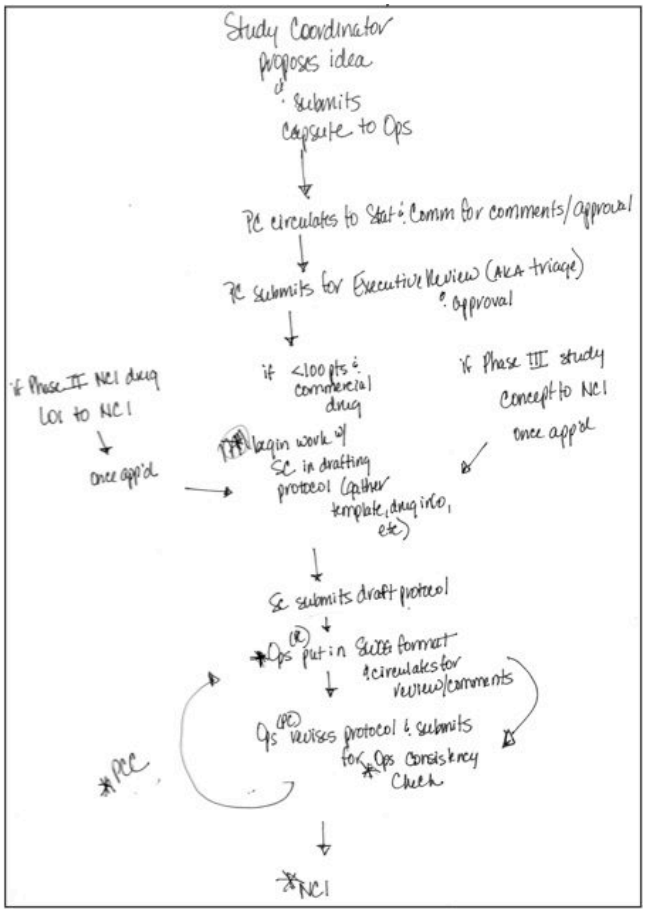

Figure 5. Process as viewed by an experienced coordinator.

4.4 Different Processes – Coordinated Work

SWOG regards clinical trial protocol authoring as a collaborative effort. Organizationally, it values active participation of the primary contributors, the PI, the PC, and stats group, but takes very seriously other individual and group contributions through PRC and CTEP meetings. The interviewees, as well, consider the protocol authoring process collaborative. In many cases the PIs list SWOG stats group as a source of critical input. PCs consider the PI and peer review committees as having important input. As we have shown, these participants maintain different representations of the authoring process. On some level, the participants follow different work processes.

In considering the participants different representations of the process two key questions arise; where do all these process representations come from and are they useful to the participants of the collaboration? The processes of collaborative authoring described here are all representations of a plan for authoring a clinical trial protocol [19]. And there are probably many more representations as well.

That the participants do not share complete understanding of the process is not surprising. From a situated action [19] perspective individuals come to some shared understanding, but that understanding is bound in one specific instance of inter-

3. The term "Study Coordinator" or "SC" is SWOG's official term for the Principle Investigator (PI) on a protocol. The term PI is used colloquially in the way we have used it here.

action. However, over the many instances of collaboration and shared understanding why don't the processes share more similarity? The breakdowns that require negotiation might be part of the exceptional nature of work, but those exceptions are often glossed over or omitted when participants are asked about the process. So, based on many collaborations with many people over time, participating in a similar (or the same) process, why don't the rational representations converge?

Another possible explanation is that maybe there is no need for the process views to converge. Weick [20] describes a model for group formation that is different than most (c.f., pp. 90-97). Common models of group formation hold that members of a group will often decide on 'what' (the ends, the tasks, the goals) of the group before settling on the 'how' (the means, the processes). Many theories of group formation adhere to this group formation model (c.f. [8]). Weick inverts this order and suggests that groups first converge on the 'how' before they settle exactly on 'what.'

One reason for the partially overlapping views of the process is that each individual aligns the process with his or her organization's goals. In our study, individuals participate in the collaborative process of creating a clinical trial protocol, but with slightly different goals. Statisticians are interested in high-quality protocols with clean statistical designs because in the end they have to analyze the resulting data. PIs participate to do medical research and improve health care – so the emphasis is on the research topic, not the authoring process. PIs want to improve the state of health care, and make a significant research contribution. PRC and CTEP committees are often focused on safety, clarity, and matching the standards in the protocol to as many local IRBs as possible. PCs are employees of SWOG and are interested in the effectiveness of the authoring process. Indeed, they all participate with different goals, but still accomplish the work of writing a protocol.

Weick suggests that what makes this all possible is mutual equivalence with minimal knowledge sharing. Weick [20] defines a mutual equivalence structure to be when one participant in a collaboration is dependent on the output of another. The output could be almost anything, Weick talks about mutual equivalence for conversations as well as knowledge work. One advantage of mutual equivalence is that individuals do not have to share the same goals to be able to sustain collaboration; they only need to agree on some general parameters for the specific instance.

In the case of protocol authoring, the participants must have some minimal understanding of the mechanisms they will use to exchange information, exchange drafts, communicate and coordinate their activities. A small number of tools that facilitate distance collaboration are assumed (e.g., email with attachments, Word, phone and voice mail) and everyone is assumed to know these tools. As well, the PC works as a type of communication and coordination hub – though at some perceived personal costs in effectiveness.

This general perspective seems to fit our data. In addition, it fits with the wide variety of backgrounds and domains of expertise of the participants. For example, the PIs are viewed as higher in status than the PCs, while the biostatiticians provide a sort of expertise that is orthogonal to the PIs knowledge. Given this disparity of backgrounds and status, perhaps it should not be surprising that different participants have different views or representations of the process. While PCs work as a communication and coordination hub, the tools allow them to make limited demands on their asynchronous collaborators. The process, timeline for milestones, and intermediary products are often negotiated for each protocol.

The data we have presented so far does not highlight specific breakdowns; instances where action is negotiated. From our involvement in the process and from studying PRC comments, we observed that when breakdowns happen they are resolved through communication. Individuals in the process can have different frames of 'where we are' and 'where we are going' in the process while still collaborating over specific details. Given PRC comments, a PC and PI will engage in negotiation about specific aspects of the collaborative writing. This is not simply clarifying a specific comment. The negotiations surround how to proceed, and what are the next immediate tasks given a PRC comment.

It seems clear from our data that individuals in collaboration can effectively coordinate their activity while maintaining only partially consistent or partially overlapping representations of what they are doing. Any single action can be negotiated, but negotiating a single instance will not necessarily align the overall trajectory of two or more individuals' conceptions of the process. Our findings suggest that support for multiple views of not just a developing document but the process of creation is a critical aspect of collaborative writing systems. The majority of collaborative writing systems and all current commercial word processors focus on the views of the document, and not onexplicit support of the collaborative process.

5. RECONSIDERING COLLABORATIVE AUTHORING

SWOG managers believe the protocol authoring 'process' can be improved. The organization is chartered to produce the three top rated protocols for each disease site (e.g., prostate cancer, breast cancer, lung cancer, etc.) in a 6 month period. The reality is that only a few protocols meet or beat the 6 month production schedule. SWOG managers have a general belief that better control over the authoring 'process' can be directly translated into more efficient production of clinical trial protocol documents.

Since people working with SWOG belong to different organizations, they usually have work responsibilities outside that of protocol development. For participants who are not SWOG employees, such as PIs or drug company representatives, protocol writing is not their top priority. In this setting participants tend to be polite. If someone asks for input from another person and if that person does not respond, the first person may wait a significant time before making the request again.

However, because the collaboration of protocol authoring spans different organizations, the SWOG management perception that the problems can be resolved by better control over the process alone is probably a mirage. The loose coupling of the collaboration, the plausibility that another member in the collaboration will produce something needed in a reasonable time, facilitates the ultimate creation of the document. The different participants assume that what they produce will be used and that they will be informed when the next contribution is required.

The organization invited us to perform the research to help understand and improve the 'process' of protocol development. Our exploration helps understand how different partici-

pants view the authoring process, and also helps frame better questions about the problems in this type of collaborative authoring. The interviewees talk about their problems with the process relative to existing tools, problems managing document complexity, and communication breakdowns. Our current recommendations are not as much about the process itself; based on Weick's view of mutual equivalence, differing views can facilitate asynchronous collaboration. Nonetheless, SWOG and our research results identify problems that can be ameliorated through the development of new tools or the extension of existing tools. We highlight some problems (awareness and coordination, document management, and communication breakdowns) that will be addressed in our continuing work with SWOG. But first we highlight changes in tools and tool use compared to prior research results.

5.1 Tools

In some ways, things are little changed since Beck & Bellotti [2] conducted their study. One problem they noted was problems with compatible editors – this is largely not a problem now as everyone uses Microsoft Word. But that results in other problems because Word is designed as a single user application [6]. For example, SWOG has attempted to use "track changes" features. However, for a number of reasons, these features do not provide sufficient support for collaborative work. Track changes can be turned off and on by users, so changes can be (accidentally) hidden. Also, Word provides insufficient support for commenting; one cannot address a comment to a specific person, nor can one make a comment on a comment, nor can one easily integrate comments from different users into a single draft. Finally, Word does not provide any support for work coordination. For all of these reasons, SWOG abandoned the use of these Word features to overcome the collaborative complexity they face.

SWOG has made some attempt to improve document management and sharing. SWOG attempted to build custom software that would help manage documents and revisions on a central server. Although a prototype system was constructed, it was not adopted by the organization, due to user-interface complexities—use of the system was perceived as more difficult and time-consuming than existing work practices. However, developers both at SWOG and at the national level in CTEP, are still aiming to build a custom-tailored groupware application for this domain. The ubiquity of Word has made one requirement clear; whatever tool is built must support reasonably seamless import/export to/from Word.

The use of email as a collaborative authoring tool has changed dramatically since Beck & Bellotti [2]. At SWOG email is the primary tool for facilitating document sharing and coordinating activity. Email is well suited to some aspects of protocol authoring: it allows collaborators to work at a distance and to work asynchronously. However, it is not very effective for conflict resolution when working through different points of view. As we describe below, email conversations and attachments can lead to awareness and coordination problems, common document management problems, and communication breakdowns.

5.2 Awareness and Coordination Activity

As we mentioned the SWOG Protocol Coordinator (PC) is a central figure in coordinating much of the on-going activity of protocol development and authoring. Email has made much of this possible. PCs like email because it avoids problems of scheduling phone calls with already overly busy PIs and the alternative, phone tag. PIs like email because they can use small amounts of unscheduled time to make progress on outside projects (like a clinical trial protocol). But this level of asynchrony clearly has costs.

One problem in using email to support these types of asynchronous and distributed processes is that it can be hard for participants be aware of each other's activity. For example, in the following email, a PC highlights how they don't know what's going on with one of the participants.[4]

> *Hello Alice, Just an update, I had sent the draft to George and Yolanda as you requested. I have heard back from Yolanda, but not George. I don't know what the situation with George is, I'm going to go ahead and incorporate Yolanda's comments and send you a draft back within the next couple of days. Did Dave have any further comments regarding the draft? Thanks, Brian.*

A PC, a SWOG employee who wants to keep the protocol process moving, will sometimes make a decision (like in the email above) to incorporate comments by one person, only to find out that another participant has made conflicting comments to what is now an older draft. One explanation for this sort of coordination problem is that participants may have only partially overlapping views of the process. The PC views the process as one of iterative development where his role is to shepherd the work forward. In contrast, it may be that George (a PI) views his participation as optional or less important at this stage in the authoring.

There are no simple solutions to the problems of awareness and coordination for the PCs at SWOG. Email has enabled their current asynchronous, geographically distributed authoring. Email was never meant to provide the level of awareness and coordination that the PCs would like. Knowledge workers now 'live' in email, and this is leading to the development of more sophisticated email environments that might help solve some of these problems. None of these email tools has emerged from the research laboratory for adoption and study. The problems with awareness and coordination are one driving force behind SWOG's desire for tools to support protocol authoring.

5.3 Document Management Problems

SWOGs document management practices, designed to simplify complexity, can actually result in problems. In the current practice the PC is a 'single scriber' [14] incorporating all changes to the document before declaring a new version. PCs have had trouble managing all of the drafts, so they developed a practice of only keeping the current draft in an electronic form. All old drafts are printed and kept on paper, but deleted from the computer. If a PC needs to retrieve information from previous versions, it is cumbersome to search paper documents and then copy information back into the current electronic version.

The PCs make a distinction between a 'draft' and a 'version' which is useful. A version is generally the last draft produced before some milestone, such as a review meeting. Versions are currently defined by their form, PDF. This is used as a signal that the current version is not open to any more modifications until after the review has been completed. Once a new version is created, the PC will attach the draft to an email that explains

4. In the following data, participants' names are anonymized.

the changes that have been made and the sections of the document which are incomplete The problem is that much of the meta data related to a given version is buried in email, and not stored with any version or draft.

Protocol review is an important step in the protocol development process to ensure the quality of a protocol document. Reviewers typically print a protocol document, write comments on a separate sheet, and then email the comments to the PC. This creates a context problem, so reviewers often write extra text to specify the protocol context (where in the document to apply the comment). For example, to make a change of only four words one reviewer wrote:

> *... sec 2, second paragraph, first sentence, suggest 'over the use of single agents' instead of 'over the sequential use of them' since we decided against a crossover.*

The PC will cross check the comments relative to the reviewed version of the document. In some cases, reviewers will supply conflicting suggestions for exactly the same sentence. Based on a review the PC will begin a range of activity to address the questions and issues raised. The PC will manually group comments that are addressing the same problems together to detect conflicting opinions. In these cases, the PC must resolve the conflict between the parties, often through email.

Document management is not just a problem for PCs, it also is problematic for PIs, statisticians and SWOG managers who follow the protocol through its development. We illustrate this with an email exchange that took place during a period of intense activity that lasted about 4 hours. A nearly complete protocol was being finished so that it could be submitted to CTEP. In preparation, the protocol was sent to Daniel, a physician, for informal review. Daniel commented on the protocol and returned it as an email attachment to all the participants, PIs, statisticians, the PC and relevant managers.

After about two hours, Jason, one of the PIs, sends another draft in response to Daniel's comments to all parties.

> *this is some additional minor changes from Daniel's. It'll help to have this cleaned up, since it's hard to track many levels of changes. -Jason*

About 30 minutes later Jason sends another email to all parties, but specifically addresses the PC. He attaches another protocol draft.

> *Korine, I concur with just about every suggested edit, and I just made a few tiny additional changes. Can you please clean this up without the cross-outs and resend, so that we can review this in a prefinal version before anticipated send-off tomorrow? References also need to be placed in a uniform format. Many thanks. - Jason*

Shortly after Jason's second email, the other PI working on the protocol sends out an email to all parties, trying to figure out which draft is currently the most up to date.

> *which version is this??? Daniel's or your edits of Daniels? I WILL review, but only the latest version. ...*

Shortly after the second PI's email, we received email from a SWOG manager, who had been following the development of this protocol: "*John, just FYI as an example of what goes wrong.*"

Document management is a tough problem with no simple solutions. But in this setting there are a number of possible changes that could help improve the process of collaborative authoring. One thing is to work with the PCs to help them manage old drafts. The problems the PCs have with their current tools results from existing practices, such as cc:ing every possibly relevant person on a protocol. This cannot be completely resolved. However, providing a backing store for old drafts might help lower the complexity of their current workspace (the reason for deleting old drafts).

The problems that PCs have with managing reviews are a different type of document management problem. On-line reviewing systems exist, and adopting one might relieve some of the current problems. But for the PC there will still be the problem of integrating the results of addressing review comments into a new version of the protocol. This problem highlights the subtle distinction between many of the current reviewing systems and the ideals of many collaborative authoring systems.

5.4 Communication Breakdowns

The participants in the collaborative authoring process rely on email as the primary communication medium. The types of communication breakdowns which the participants experience are a function of common problems with email and email readers.

Email is poor at representing the threading of conversations, especially those involving multiple participants. In the authoring process, there are lots of conversations and discussions. This classic problem makes it hard to track the topic and idea flow as a document develops. The use of email attachments results in extra work to explain the changes in a draft. When a participant responds to a request by a PC with an attachment, it requires the PC to track which draft was used.

A statistician usually gives revision comments in email. Such comments may apply to multiple places in the protocol draft. For hard-to-explain changes the change will be made in the draft and a simple explanation will be provided in email. In some cases, changes will be used to initiate discussion with other collaborators on the protocol. Tracking these kinds of communications and the resulting changes to the protocol is one job of the PC. Tracking these events is even more difficult when the PC is not included in the email exchange.

Some communication breakdowns relate to the roles that the participants take in the protocol writing process. As mentioned before, the PC generally sets the pace of the collaboration. The PC is also a hub for most of the communication. Often a PC has to relay email with questions and answers to those involved. Communication breakdowns can result from this indirect communication process. Some questions are asked repeatedly. Some answers conflict with each other. A complaint that we heard many times during our interviews is that many statisticians do not know when or if their questions have been addressed. To the statistician, these questions are viewed as very important for statistical validity, while other participants in the process (possibly including the PI) may not share this view. Once again, different participants only partially share views of the overall authoring process.

While we do not explicitly consider a delay to be a communication breakdown, excessive delay is problematic. In some cases a small delay can be good in that the person responding can have time to formulate a reasonable response. The general problem that results in delay is that participants do not have a good idea about the changes that occur in each other's work. For example, when a PI's (a practicing academic physician)

schedule changes and she is required to spend more time on a clinical ward, she will typically take longer to respond.

Improving communication in this situation is not just a function of switching tools nor encouraging more communication; the PCs have enough communication to manage already. Because writing a clinical trial protocol can take more than 6 months participants' schedules will change. This is a space where awareness tools could augment the existing practices to help clarify, perhaps, why a response is delayed. Participants might be able, then, to create other mechanisms to encourage a timely response.

6. SUMMARY

Collaborative authoring and group editing systems have been studied and developed for more than a decade. Despite improvements in collaboration tools, file sharing and email application users still struggle with collaborative authoring; collaborative authoring is still not well understood.

Our data demonstrates that the participants do **not** share the same view or representation of the authoring process, even though it has a long organizational history. However, the participants still use their partially consistent representations to coordinate their work. This partial consistency is a feature that facilitates asynchronous collaboration at a distance. Based on the results we identified possible improvements in general tool use, document management practices, and communication. Our ongoing work with SWOG includes tailoring existing tools and the design and development of a new tool [21] to improve the collaborative authoring of clinical trial protocols.

ACKNOWLEDGMENTS

The authors recognize the many employees of SWOG who have generously given their time to answer our questions. In particular, John Crowley, Director of the SWOG Statistical center has encouraged us to continue this work. As well, we are grateful for the financial support SWOG has provided to the second author. Lastly, we thank the anonymous reviewers who provided many insightful comments and Kjeld Schmidt who took the time to help us with our revisions.

7. REFERENCES

1. Beck, E.E. A Survey of Experiences of Collaborative Writing. in Sharples, M. ed. *Computer Supported Collaborative Writing*, Springer Verlag, 1993, 9-28.
2. Beck, E.E. and Bellotti, V.M.E., Informed Opportunism as Strategy: Supporting Coordination in Distributed Collaborative Writing. in *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work (ECSCW '93)*, (1993), 219-231.
3. Dillon, A. How Collaborative is Collaborative Writing? An Analysis of the Production of Two Technical Reports. in Sharples, M. ed., Springer Verlag, London, 1993, 69-86.
4. Fazi, P., Luzi, D., Manco, M., Ricci, F.L., Toffoli, G. and Vignetti, M., WITH: A System to Write Clinical Trials using XML with RDBMS. in *Proceedings of the 2002 American Medical Informatics Association Annual Fall Symposium (AMIA'02)*, (2002), 240-244.
5. Flower, L.S. and Hayes, J.R. The Dynamics of Composing: Making Plans and Juggling Constraints. in Steinberg, L.G.a.E. ed. *Cognitive Processes in Writing: an Interdisciplinary Approach*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1980, 31-49.
6. Grudin, J. Eight Challenges for Developers. *Communications of the ACM*, *37* (1). 93-104.
7. Kirby, A. and Rodden, T., Contact: Support for Distributed Cooperative Writing. in *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work (ECSCW'95)*, (1995), 101-116.
8. McGrath, J.E. *Groups: Interaction and Performance*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
9. Mitchell, A., Posner, I.R. and Baecker, R.M., Learning to Write Together Using Groupware. in *Proceedings of the 1995 ACM Conference on Human Factors in Computing Systems (CHI'95)*, (1995), 288-295.
10. Modgil, S. and Hammond, P. Decision Support Tools for Clinical Trial Design. *Artificial Intelligence in Medicine*, *27*. 181-200.
11. Neuwirth, C.M., Kaufer, D.S., Chandhok, R. and Morris, J.H., Computer Support for Distributed Collaborative Writing: Defining Parameters of Interaction. *Proceedings of the 1994 ACM Conference on Computer-Supported Cooperative Work (CSCW '94)*, (1994), 145-152.
12. Olson, J., Olson, G.M., Storrosten, M. and Carter, M., How a Group-Editor Changes the Character of a Design Meeting as well as its Outcome. in *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work (CSCW'92)*, (1992), 91-98.
13. Pattow, D. and Wresch, W. *Communicating Technical Information: A Guide for the Electronic Age*. Prentice Hall, Englewood, NJ, 1993.
14. Posner, I.R. and Baecker, R.M., How People Write Together. in *Proceedings of Twenty-Fifth Hawaii International Conference on the System Sciences (HICSS)*, (Hawaii, 1992), IEEE.
15. Shahar, Y., Shalom, E., Mayaffit, A., Young, O., Galperin, M., Martins, S. and Goldstein, M., A Distributed, Collaborative, Structuring Model for a Clinical Guideline Digital Library. in *Proceedings of the 2003 American Medical Informatics Association Annual Fall Symposium (AMIA'03)*, (2003).
16. Sharples, M. An Account of Writing as Creative Design. in Levy, C.M. and Ransdell, S. eds. *The Science of Writing: Theories, Methods, Individual Differences, and Applications*, Lawrence Erlbaum Associates, 1996.
17. Sim, I., Olasov, B. and Carini, S., The Trial Bank System: Capturing Randomized Trials for Evidence-Based Medicine. in *Proceedings of the 2003 American Medical Informatics Association Annual Fall Symposium (AMIA'03)*, (2003).
18. Suchman, L.A. Making Work Visible. *Communications of the ACM*, *38* (9). 56-64.
19. Suchman, L.A. *Plans and Situation Actions: The Problem of Human-Machine Communication*. Cambridge University Press, Cambridge, 1987.
20. Weick, K.E. *The Social Psychology of Organizing*. Random House, New York, 1979.
21. Weng, C., McDonald, D.W. and Gennari, J.H., A Collaborative Clinical Trial Protocol Writing System. in *Proceedings of the 11th World Congress on Medical Informatics (MedInfo'04)*, (San Francisco, 2004).