# Evaluating Computer-Supported Cooperative Work: Models and Frameworks

Dennis C. Neale
Center for Human-Computer Interaction
Department of Computer Science
Virginia Tech, Blacksburg, VA 24061
+1 540 231 7542
dneale@vt.edu

John M. Carroll & Mary Beth Rosson
School of Information Sciences and Technology
The Pennsylvania State University
University Park, PA 16802
+1 814 863 2476
jcarroll@ist.psu.edu, mrosson@ist.psu.edu

## ABSTRACT

Evaluating distributed CSCW applications is a difficult endeavor. Frameworks and methodologies for structuring this type of evaluation have become a central concern for CSCW researchers. In this paper we describe the problems involved in evaluating remote collaborations, and we review some of the more prominent conceptual frameworks of group interaction that have driven CSCW evaluation in the past. A multifaceted evaluation framework is presented that approaches the problem from the relationships underlying joint awareness, communication, collaboration, coordination, and work coupling. Finally, recommendations for carrying out multifaceted evaluations of remote interaction are provided.

**Categories & Subject Descriptors:** H.4.1 [**Information Systems Applications**]: Office Automation - groupware

**General Terms:** Measurement, Theory, Human Factors

**Keywords:** CSCW evaluation, models, awareness, common ground

## 1. INTRODUCTION

Evaluation approaches to computer technologies must adapt and change significantly if we are to keep pace with evolving human-computer interaction. Much evaluation of the past was concerned with the cognitive functioning of a single user sitting alone in front of a computer display. Users were modeled as vigilant, task-oriented workers operating in relatively narrow contexts over short time periods without regard to their broader functioning as social members of larger groups and communities. Computer-supported cooperative work (CSCW) evaluation has been, in general, more broad in nature, but it often has been ill defined, time consuming, labor intensive, difficult to implement, difficult to interpret, and largely ineffective at producing timely formative data that is needed if groupware applications are to succeed.

New evaluation strategies are needed that uncover central issues associated with groupware success and failure, and they need to be more flexible than they currently are in order to adapt to a greater range of factors that need to be considered. This paper will

begin by describing several key challenges to distributed CSCW evaluation. In particular this work focuses on evaluation strategies for remote collaboration involving long-term activities that include both synchronous and asynchronous interaction. The types of complex tasks and teams associated with this type of interaction will be described. A review of the current state of evaluation approaches and methods in CSCW will be outlined. A new model will be provided that approaches evaluation by targeting the underlying processes of human collaboration in groupware systems. This approach is complementary to other conceptual frameworks and methods. Lastly, design recommendations for carrying out multifaceted evaluation approaches will be outlined.

## 2. REMOTE EVALUATION IS DIFFICULT

Why has the evaluation of CSCW systems, in particular systems that support remote collaboration, failed to satisfy the demands? The quantitative metrics used to evaluate user interfaces in many cases has been elusive. In fact, having limited quantitative performance data that does not measure the critical criteria that defines CSCW failure or success can be useless or worse misleading. In the case of collaborative applications, performance measures alone rarely are good indicators for improving CSCW systems. On the other side of the continuum, many workplace evaluation approaches have ignored quantitative data, relying entirely on naturalistic inquiry and description. These approaches in isolation also often fail to produce useful criteria for designing CSCW systems, including the lack of generalizability to different contexts.

The evaluation of distributed CSCW systems has been too frequently method driven by various disciplinary preferences, rather than driven by frameworks that get the appropriate questions answered. This has been particularly true for distributed systems that support remote collaboration involving complex synchronous and asynchronous interaction where the central underlying variables have not been fully identified or understood. This paper focuses on these types of systems and does not necessarily reflect entirely on all types of CSCW evaluation. The first step in solving the evaluation crisis facing distributed system development is to understand what the barriers are in evaluating these types of systems. Three problems have made this type of evaluation difficult: (1) The logistics in carrying out distributed evaluation are difficult; (2) There are a greater number of variables to consider, and they are more complex; and (3) evaluation in much of CSCW needs to focus on validating the re-engineering of group work based on CSCW concepts.

## 2.1 Logistics of Data Collection

Activities that must be evaluated are often distributed in time and place. The pragmatics of negotiating data collection is difficult under these conditions. When, where, and how to collect data can be a significant problem facing researchers. Multiple evaluators are needed to capture distributed interaction. Much of the interaction of interest occurs at times that are either inaccessible to evaluators or occur over long time periods making it impractical to capture. Even when it is possible to collect data, it can be difficult to predict when and where the interaction of most interest is going to occur. Having only "snapshots" of the total interaction that is relevant leaves evaluators wondering whether more data is needed. All these factors make it difficult to prioritize the most appropriate data collection strategies.

## 2.2 Number and Complexity of Variables

Evaluating CSCW systems is difficult, much more difficult than single-user systems of the past. The variables that need to be considered are more diverse and complex. Individual cognitive factors must be considered, as well as cooperative and collaborative factors, usability issues for individuals and groups (ease of use, effectiveness, efficiency, satisfaction), the social and organizational impact, and the larger context that situates the other factors. This makes it difficult to know where to begin. The underlying causes to problems are often distributed in time and space [28]. Multiple factors often contribute to any given measurement or subsequent interpretation. In fact, these factors are typically not fully proximal to the situation where observations are made. This makes the constructs under consideration much more inaccessible than what is available to the researcher evaluating real-time, dynamic systems. This also makes it more difficult to translate behavioral findings from the evaluation into system requirements and design solutions and outcomes.

## 2.3 Validating Re-Engineered Group Work

Much of CSCW evaluation must occur with relatively developed systems in real contexts [12]. One important reason for doing this is that introducing cooperative tools alters the group interaction, and therefore, it is the only way to understand basic characteristics of the teamwork that will ultimately result when collaborative tools are adopted. It is natural for researchers to want to focus on outcome measures at this stage, and there is precedence for this approach in HCI. However, process measures may be more important for determining outcomes well into the general use of collaborative tools by groups. Generally speaking, more complex CSCW applications have not moved beyond exploratory systems, and the prerequisite understanding of groups and organizations that is required to mature these systems is lacking [30]. This makes it difficult to get initial designs even close the first couple of design cycles. As a result, there are more iterations of the requirements, design, and test phases that take longer than with single-user systems.

Given this situation, CSCW evaluation must focus on the validation of CSCW concepts well into their general adoption. The emphasis must continually be placed on issues at the group and organization level. Focusing on lower-level behavioral data may be inappropriate if problems at the social and organizational levels are not addressed [17]. The re-engineering of work systems and its consequences on group outcomes needs to be a central priority in CSCW evaluation. This is a paradigm shift toward socially centered design from past design eras of system-centered and user-centered design [42].

## 3. LONG-TERM ACTIVITIES AND TEAMS

The types of evaluation being referred to in this paper are intended to address complex systems that are distributed across locations and support synchronous and asynchronous interaction. This refers to software systems that support goal-oriented teams working in collaboration to carry out joint projects that are characterized by the need for communication, planning, coordinating tasks, monitoring project progress, and cooperation. Ill-structured group processes characterize such activities. These types of systems (and the activities they support) and the teams who use them have unique properties that should be a central consideration in this type of evaluation.

## 3.1 Long-Term Activities

The long-term temporal structure of activity mediated by collaborative technologies has significant consequences for groups [23]. The project work associated with these systems involves long-term activities ranging from weeks to years where people must establish and maintain an ongoing awareness of other's actions, plans, goals, and activities. These types of activities are goal-oriented and involve planning, acting, assessment, and adaptation to iterative re-planning based on changing objectives and circumstances. Inherent in these types of groups is the need for information sharing, scheduling, role taking, synchronization, and allocation of resources. Evaluation must consider the sequential and longitudinal characteristics of long-term activities. Teams of people often are engaged in these types of long-term activities.

## 3.2 Teams

Groups are not the same as teams. Groups have task structures with limited role differentiation, and performance depends largely on individual efforts. Teams, on the other hand, have members with specialized roles, and the team works together to accomplish common goals [14]. The interdependence of tasks and their coordination are defining characteristics of a team. The consequences for evaluating teams are significant. Evaluation is more complex for teams than groups because although the central factors are the same – communication, coordination, cooperation, awareness – for teams it is the aggregate of these factors that must be considered. Much of the prerequisite research used to understand CSCW has been based on groups rather than teams. Loosely formed groups have always existed. However, teamwork is becoming more prevalent today because the complexity of work has increased, demanding that more tightly coupled groups of people carryout tasks with a common goal. Teams are more frequently faced with distributed group participation as well. The work described here focuses on teams that are task-oriented and operate in distributed organizational settings.

Several evaluation frameworks have begun to address these issues, some more successfully than others. However, there is still further work needed for developing evaluation models that more specifically address complex research problems for long-term distributed team collaboration. In the next section, several of the prominent CSCW evaluation frameworks are reviewed.

## 4. CSCW EVALUATION

One of the most difficult problems facing CSCW evaluation is working out interdisciplinary differences. Disparity in approaches

is often the result of the diverse backgrounds of researchers and disciplines involved in studying groupware systems. Two broad categories that characterize different approaches to evaluation are the quantitative and qualitative paradigms. The quantitative paradigm is referred to as the empiricist or positivist perspective. The qualitative approach is described as constructivist or naturalistic. Many debates have raged in the social and behavioral sciences over the "nature of reality" and how to measure it based on these two approaches. The dominant methodological paradigm of group research has been from the positivist perspective [1]. This research has served as the underlying knowledge base for CSCW systems.

The direct study of CSCW systems has been a mixed effort. Workplace studies have been most associated with ethnography, but many concerns have been raised regarding its role in CSCW design [37]. However, much of this work in CSCW is really qualitative research [25], rather than true ethnographic studies. There are a greater number of laboratory experiments reported in the literature [34], and there is a rich tradition for this approach in engineering and computer science. But laboratory studies have been criticized as well for being ineffective as a paradigm for evaluating CSCW [12,1]. The frameworks reviewed below have varying degrees of association with both the qualitative and quantitative paradigms.

Evaluation frameworks in the literature fall into three different camps. *Methodology-oriented frameworks* describe the types of experiments and methodologies available to CSCW researchers. These frameworks are useful for understanding the general types of evaluation possible, but they provide little guidance for choosing among different types of methods. Other frameworks have described the group factors that should be considered during evaluation. These are *conceptual CSCW frameworks* for discerning what should be evaluated in CSCW. However, there is little literature mapping conceptual constructs to methodological approaches. Even if evaluators know what factors are important and have a variety of methods available to them, it is difficult to determine which methods will yield the most effective findings from the various issues under consideration. Lastly, *concept-oriented frameworks* focus on specific aspects of group behaviors or concepts, such as communication or coordination. These frameworks are more limited, but they do offer specific advice for focusing on limited or isolated aspects of group interaction. It is, however, difficult to determine how methods can be combined to form comprehensive approaches. There has been some effort to unify these differing perspectives [45], but much more work is needed. Understanding the levels of analysis in CSCW provides a basis for appreciating how the different frameworks play a role in evaluation.

## 4.1 Levels of Analysis in CSCW

Several different levels of evaluation and analysis are possible in CSCW. The individual, group (team), organization, and industry are common levels of analysis for CSCW systems [30]. Individual, group, and organizational levels map relatively well to cognitive, rational, and social bands of human activity. Researchers often emphasize different event types, and this dictates what behavioral data gets collected. Most evaluation in HCI has focused on short-duration tasks, such as those typically examined in usability studies. However, the real challenge in evaluating future systems will be to consider long-term use settings where behaviors occur over months, years, and decades.

Many evaluation methods, frameworks, and analyses stem directly from theory related to one of the levels. Theory plays an important role, often explicitly and implicitly, in what data collection techniques get used, how the data is analyzed, and ultimately how the findings are interpreted. Many of the methods underlying evaluation models are biased toward a particular perspective. It is important to keep in mind the level of analysis and the methods used as we consider the appropriateness of the evaluation frameworks reviewed.

## 4.2 Existing CSCW Evaluation Frameworks

There are countless books on research methods, but McGrath's reviews of research strategies appropriate for groups are well known to many CSCW researchers [24, 22]. McGrath outlines a number of research strategies available to the CSCW researcher, and he examines some common types of measures that correspond to different research methods. He points out that researchers are always trying to maximize generalizability, precision, and realism, but it cannot be done simultaneously with a single approach or method. Multiple methods must be used to balance shortcomings with any given approach. Other examples provide more limited, but similar taxonomies [46, 34]. Although these frameworks are helpful for understanding the types of evaluation possible, in general, they do little to help the researcher map methods to the constructs of interest in CSCW. Others have provided conceptual frameworks of group behavior for understanding what should be evaluated.

A number of conceptual frameworks have been proposed that outline the major factors relevant to analyzing CSCW [22, 36, 32]. They have several properties in common: Group characteristics, situation factors (context), individual characteristics, task properties, group process, and task and group outcomes. Each one of these factors can have a number of issues associated with them. Much of these frameworks stem from early research on group behavior. The factors in these frameworks correspond generally to situation, task, and human considerations in any type of applied research endeavor. However, there are other frameworks with different approaches [40, 30]. Pinelle, Gutwin, and Greenberg [35] have recently developed a framework for conducting groupware usability evaluations that focus on the mechanics of collaboration. Aside from gross descriptions like laboratory or field studies, it is difficult to determine how these factors should be studied, and it is difficult to determine what methods are best suited to which factors. However, there are a large number of research papers that describe various evaluation measures for specific types of circumstances.

Concept-oriented frameworks describe how specific methods can be used to measure concepts like communication effectiveness, awareness, or trust. For example, video analysis methods have been described for multiple sites [39], participatory design methods for groups [38], data logging methods for multi-user applications [16], and several methods (e.g., activity set analysis) for measuring interpersonal awareness [49]. Breakdown analysis is another more general method for studying how groups encounter problems [15]. These sources are useful for understanding how to implement specific methods, but it is difficult to situate any given method in the larger evaluation approach or to understand how to come up with a comprehensive set of measures for addressing all of the constructs of interest to the evaluator. In the following sections, we first describe the concept of activity awareness and then present a new model for evaluating distributed collaboration based on awareness.

## 5. ACTIVITY AWARENESS

Perhaps the core challenge for CSCW systems is providing effective support for activity awareness. Collaborators who are not able to be at the same place at the same time need continuing support to remain aware of the presence of their counterparts, their tools and other resources, their knowledge and expectations, their persistent attitudes and current goals, the criteria they will use to evaluate joint outcomes, and the current focus of their attention and action.

Many concepts of awareness have been discussed in CSCW literature: social awareness, presence awareness, action awareness, workspace awareness, situation awareness. This variety is itself an indication of the importance of awareness to CSCW designs, and of course to the experiences of users. But it also suggests that a more encompassing concept is required. We suggest the term "activity awareness", incorporating the term activity from the very broad and muti-layered concept from activity theory.

Activities are substantial and coherent endeavors directed at meaningful objectives like "designing the layout of a town park". Longer term activity entails goal decomposition, nonlinear development of partially-ordered plan fragments, interleaving of planning, acting, and evaluation, and opportunistic plan revision. It involves coordinating and carrying out different types of task components, such as assigning roles, making decisions, negotiating, prioritizing, and so forth. These components must be understood and pursued in the context of the overall purpose of a shared activity, the goals and requirements for completing it, and how individual tasks fit into the group's overall plan. To more fully understand the role activity awareness plays in remote collaboration, we have developed a new model of awareness evaluation.

## 6. AWARENESS EVALUATION MODEL

As already reviewed, there are several conceptual frameworks that structure the important variables to consider when evaluating CSCW applications. The framework we present here targets distributed applications. Other alternative interpretations based on different perspectives can be equally valid. This framework is not intended to be a theory about distributed interaction. Its purpose is to provide a model or map of the important variables to consider during evaluation. In this sense it is a conceptual CSCW framework for distributed applications. One of the goals of this model is to simplify the important factors for understanding the relationships between variables. Not all aspects of a complex system can be evaluated at once. Understanding the relationships between variables can help in prioritizing the most important factors to study given varying requirements and circumstances.

Figure 1 shows the major variables considered in the awareness evaluation model. Contextual factors underlie all collaborative activities and shape how the work is structured. Work can be loosely or tightly coupled based on the communication demands of the activities. More tightly coupled work requires greater demands on the communication. The greater the work coupling, the greater the demand for coordinated behaviors as well. Distributed process loss results from the amount of coordination that is required to manage the main work of interest. If the proper levels of communication and coordination are supported, groups achieve common ground and acquire activity awareness critical for effective group functioning. However, increases in these same factors places demands for greater common ground and awareness.

This model focuses on the central relationships underlying the processes of distributed group work. Communication, coordination, and work coupling form the basis for explaining how successful groups will perform. And these factors are heavily constrained by contextual factors, common ground, and awareness. Each component in the framework has a number of properties that must be considered.
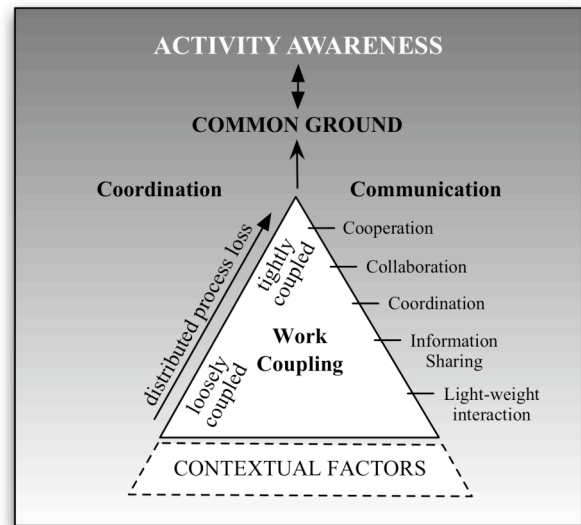


**Figure 1. Model for evaluating activity awareness**

### 6.1 Contextual Factors

The history of computer systems development shows that research has moved from physical ergonomics, to information processing at the interface, to the broader context of behavior and interaction rich with complexity. This moves the unit of analysis from human action to more comprehensive activities. When the evaluator begins to analyze interdependent activities, properties of the context become a central issue. Activities that are ongoing, spanning across people and locations, are impossible to make sense of without first understanding the context where they are situated [43].

People manage context, and the social fabric that binds it together, in very complex, yet subtle ways. People use context for understanding how to organize their individual efforts in the framework of their social interactions as part of groups. However, they do this in large part as background activity, without being fully cognizant of how context shapes their behavior. Social group dynamics are ingrained in human nature and unavailable to normal conscious inspection [13]. In normal group interactions people easily manage context because they are immersed in rich, multiple sources of information that is easily obtainable. Problems can develop in face-to-face interaction and with co-located teams, but we have an abundant set of strategies for dealing with inconsistencies in shared context in these cases.

Distributed systems fracture background contextual information significantly, especially contextual information that is temporally removed from immediate interaction. And this information is for the most part totally unsupported with current technologies. What

is even more problematic, because people manage context largely as background activity to their main activities, they behave as if they have a full understanding of the contextual variables shaping their behavior. They believe they have a common frame of reference – shared context. As a result, "surprise" breakdowns routinely develop, and their unexpected nature make them especially frustrating to users of CSCW systems. It is not just that they are unexpected; it is that they are unexpected and often directly contrary to the beliefs held by group members. In this way, external conditions that define the context shape the group's internal development, often with negative consequences. We are only beginning to understand how context shapes behavior, and therefore, are only beginning to be able to design tools that share context. It could be argued that altering shared context is the single most significant reason for the failure of distributed system use and ultimately its adoption.

What is context? Context has two parts. First, our notion of context stems from activity theory [26] – context is comprised of the activities themselves. It is more than just a container that frames activities. Context involves the internal states of the actors themselves and develops dynamically as part of normal interactions with others. Second, context is made up of the little things: Who is present? What are they doing? Are they bored? Who else is present? What is their relation to others? What are the artifacts of interest? What is going on around people? What are the subtle circumstances of peoples' lives and situations outside the immediate context shaping their behavior in the current situation? Without this information it is difficult to understand why people do what they do, especially when they do things contrary to what is expected or planned. Much of this information gets shared in lightweight, informal interactions, and it is communicated and collected through a variety of different types of information. Trying to represent it digitally is difficult, and transforming it in subtle ways may render it useless. Furthermore, making the collection of contextual information an active task for users makes the information unappealing or unusable. However, we must begin to evaluate context more seriously, but studying context to reveal the issues previously raised is a daunting task.

How should context factor into the evaluation of remote collaboration? Because most group research has come from the positivist paradigm, context is stripped away from the analytic process. Evaluation must more seriously consider the interactions between the group and embedding context [1]. Context considerations must span across individual members, the group itself, and the larger organization in which the other two are embedded. The contextual factors that span over time are often the most important.

Context for long-term, group interaction by its very nature is temporally dependent, and each component of the context (individual, group, and organization) is temporally dependent itself. Contextual information that is temporally removed from the immediate situation being evaluated often only gets shared informally and in the subtlest of ways. Interpreting what people are doing in complex groups always relies on context. Understanding what is said and done at any given moment in time in a group always exceeds what is immediately available. The more information group members have that is outside the immediate behavior, information about the global context, the more cohesive and effective the group is. This makes it especially difficult for the evaluator because it is often only possible to capture this type of information indirectly. And once it is

captured, it must be reconstructed from a variety of sources during the analytic process.

## 6.2 Work Coupling and Communication

The notion of work coupling in groups and organizations has multiple meanings across disciplines, but in CSCW it has become a concept for defining the intensity or demand of the work for information sharing or level of communication required [2, 31]. Here the notion of communication is closely intertwined with the level of interactions between group members. It is a multifaceted concept that includes aspects of the work and the demand for communication needed. The granularity of dependencies between group members for successfully completing work varies, and the degree to which members must communicate to successfully perform both relate to the degree of work coupling. Work coupling reflects the amount of individual work that can be done before one has to interact and communicate with another. Loosely coupled work requires few interactions, and the communication that does exist is effortless, uncomplicated, and straightforward. Tightly coupled work is highly dependent on frequent communication, and the communication is demanding in the sense that highly interdependent tasks depend on the quality of the communication.

Based on the literature and through the process of our own work, we have identified five levels of work coupling: Light-weight interactions, information sharing, coordination, collaboration, and cooperation. Light-weight interactions are only loosely tied to the work itself. In this case people move between causal social interaction and communication about the work. Contextual information that is not specific to communication about the work is often shared at this time. This is information about people's lives and work situations that help others contextualize behavior in the current context and across all of their interaction with the group. Information sharing can be unidirectional, or it can occur in inform-acknowledge pairs. Important background issues related to the work often arise in these exchanges, and they can make all the difference for understanding what has or will occur in relation to others.

Coordination, collaboration, and cooperation are much more tightly coupled than the previous two. Coordination requires group members to coordinate both the activities and communication. Group members must coordinate the content of the work and the process involved in carrying it out. Coordination is a significant endeavor in its own right and will be discussed in greater detail in the next section. Collaboration levels of work coupling involve group members who work toward a common goal. They often are performing separate tasks that have a high degree of interdependence, but work is still done by individual members. They share goals, tasks, and a desire to maintain a high state of shared knowledge.

Cooperation is the highest level of work coupling, and it demands the greatest amount and highest quality of communication. People at this level of work coupling have shared goals, common plans, shared tasks, and significant consultation with others about how to proceed with the work. Many of the tasks are performed face-to-face, and they are carried out concurrently as shared activities. People at this level are committed to the team efforts, and they put the team's priorities over individual goals. There is a high demand for personal contact with this level of work, and current technology does not support these kinds of activities well. Tightly coupled work is often ambiguous, ill structured, requires high

levels of problem solving, and requires constant reassessment of priorities and goals. As work moves from coordination, collaboration, to cooperation, team coordination becomes a significant aspect of the group work.

## 6.3 Coordination

Coordinating collaborative and cooperative activities is difficult, regardless of whether it is mediated by technology. Coordination is only one component of teamwork, but it is a perquisite to successful teamwork on many other levels. But what is coordination in distributed CSCW interaction? For the purposes of this discussion, coordination will be characterized in terms of processes, procedures, tasks, tools, and awareness. "Coordination is the attempt by multiple entities to act in concert in order to achieve a common goal by carrying out a script/plan they all understand." [20]. Time is a key component in this type of coordination, and simultaneity constraints include timing and dependencies between the sequences of other events [21].

Specific procedures must be in place for groups to coordinate, especially at the higher levels of work coupling. "Coordination is the set of procedures by which teams plan, organize, orchestrate and integrate their activities to achieve shared goals [50]. Procedures can be explicitly built into team interaction, or they can be implicitly used ad hoc. A number of tasks have coordination characteristics, including planning, scheduling, assembling resources, managing resources, task allocation (roles), alignment, monitoring task and activity states, information sharing, and managing interpersonal relationships. The processes and procedures for managing these coordination tasks depend heavily on the tools available to the group (e.g., shared calendars, workflow tools, whiteboards, etc.). Shared external representations of the work in distributed systems may be one of the more significant, yet under recognized coordination devices.

Coordination can be viewed as overhead, an undesirable activity that is necessary to complete other interactive group activities. The overhead or operating costs involved in coordination is referred to as process loss [41], and distributed process loss is much more costly. So costly, in fact, groups often do not recover from its effects. It can literally take over the group activities to the point where people suspend their joint activities. Brooks [3] describes how the effort increases by a factor of $n(n-1)/2$ for each task that must be separately coordinated, and this can actually counteract the collaborative efforts resulting in a net decrease in performance.

Awareness permeates these other factors. The more aware people are, the less there is a need to coordinate activities [9]. In fact, coordination can only occur if people are aware. Maintaining awareness, like coordination, however, is a background process. In addition, awareness is a mental state, and the joint awareness of group members is their common ground. In the next section we describe how the theory of common ground provides a framework for understanding distributed joint awareness and the role that the other factors play in this model of awareness evaluation.

## 6.4 Awareness and Common Ground

The joint awareness two people share is their common ground. Clark's theory of common ground maintains that people must have this shared awareness to carryout any form of joint activities [6]. Two people's common ground is the knowledge each believes the other shares in common with them. Although common ground is a general theory of language use, it holds true for all collaborative activities. To communicate, collaborate, and coordinate people must share a vast amount of information or mutual knowledge. Groups try and achieve common ground across conversational exchanges, related activities, and broader interactions that occur over long time periods. They must update their common ground on a continual bases, and they do this through a process called *grounding*. At the conversational level grounding is the continual process of trying to determine what has been said has been understood, and it is a joint effort on the part of both people in a conversation. Across multiple exchanges and extended activities, people also try and ground what they understand about each other's activities, about what artifacts they have in common, what they believe the current state of objectives and plans are, and so forth.

The properties of awareness in CSCW were described earlier. Awareness is both a process and a product. People do a great number of things to maintain awareness of other's actions and the state of shared activity and artifacts. But awareness is also a psychological concept, a mental construct or model of how aware someone is. At any given time people have some level of awareness of other people and the state of their shared world. Common ground extends the concept of awareness to the idea of a shared or joint awareness. It provides a framework for understanding how awareness functions between two people and across multiple group members. Common ground is the product of joint awareness or mutual knowledge, and grounding behaviors is the process of maintaining joint awareness.

These ideas nicely tie together other pieces in the evaluation model presented. The level of work coupling places the demands for how much communication is required, and it places the demands for the amount of grounding that is required to develop common ground. Common ground also provides a framework for understanding coordination. Coordination is the process or managerial mechanisms for completing collaborative activities, but it also serves as a process for managing common ground across group members. Coordination defines much of the mechanisms for grounding. Lastly, the amount of context group members share makes an enormous difference in the quality of their common ground. Context is the bedrock or foundation that determines the level of joint awareness or common ground people share, and it provides the catalyst for grounding mutual knowledge at all levels of human's joint activities.

## 7. USING THE AWARENESS MODEL

Over the past several years we have been developing a multifaceted evaluation framework for complex, distributed activities [27]. We began our work as part of our Learning in Networked Communities (LiNC) project [5, 19]. The LiNC project was a multi-year effort to develop and study tools to support remote collaboration in middle and high school science classrooms. We continued this work specifically targeted at studying and supporting tools for awareness [4, 10]. This work has brought to the forefront the issues described in our evaluation approach.

We developed a Java-based system called Classroom BRIDGE [10]. The system was evaluated in the local public school system over a 2-year period. Two middle school classrooms (6th and 8th grade) used the system to carryout yearlong collaborative science projects. Distributed teams made up of 2-3 students in each classroom met once a week during the entire class period (30 - 45 minutes). However, students worked independently on their projects with their partners from the same classroom at other

times during the week. The projects were carried out almost entirely using the collaborative system. We had two goals: Evaluate the role of activity awareness in distributed long-term projects, including the continual introduction of system features to support awareness, and develop and study multifaceted evaluation approaches.

Classroom BRIDGE provides students with a collaborative multimedia notebook for developing shared documents with text, graphs, tables, and images. The full editor allows real-time interaction. The entire system is also propagated to the Web, which offers more limited browser-based access to documents and others. Students primarily used an integrated chat tool for communication, but there was some limited face-to-face interaction. User lists, activity status, and location information promote synchronous awareness, and more extended activity awareness is supported through the use of an integrated calendar and timeline for planning and artifact histories. The timeline shows common deadlines and project status, as well as version histories for all documents created during the projects. A concept map interface provides a conceptual view of the relations between documents and supports document creation and organization.

A variety of data collection methods were used to evaluate student interactions: direct observation and field notes, contextual inquiry, videotaping, system logs, artifact collection, communication histories, questionnaires, and interviewing. The computer logs provided a complete record of system use, including document access and manipulation and chat communication. All synchronous interaction was captured on videotape in both locations to record proximal face-to-face communication and student-teacher interaction. Two to four researchers collected data during synchronous system use. Analysis methods included activity and work coupling sets, computer log analysis, breakdown and critical incident analysis, content analysis, statistical analysis of questionnaire data, and participatory integration of the data by research members. Data analysis was an iterative, collaborative process that involved interleaving different temporally dependent data types to reconstruct the activities distributed in time and place.

Examples of the data analysis from year one is presented here to illustrate our use of the awareness evaluation framework. At the heart of the model is the level of work coupling. Activity sets were generated from the video records categorizing behaviors into five discrete behavioral states: face-to-face interaction (with remote partners), proximal interaction (same class group members), remote interaction, focused work (little interaction with others), and parallel activities (other classroom activities). These states represented all of the students' activity during times designated for synchronous interaction. The five levels of work coupling were described earlier (light-weight interaction, information sharing, coordination, collaboration, and cooperation). A collaborative process was used to code behaviors according to the 5 levels of work coupling independent of the activity set analysis. Percent time in each activity and work-coupling level was generated.

Students spent approximately 2% of their time in face-to-face interaction, 31% in proximal, 32% in remote, 5% in focused, and 30% in parallel activities. The majority of students' time was split relatively equally among co-present work with classmates, work with remote classmates, and co-located parallel activity. In the work coupling categories, students spent approximately 14% of their time in no interaction, 25% in light-weight interaction, 13%

in information sharing, 28% in coordination, 17% in collaboration, and 3% in cooperation. To determine how the type of activity affected the level of work coupling, percent time in different levels of work coupling was calculated within each category of activity. Table 1 shows the results for percent time in each category.

**Table 1. Percent time in work coupling by activity**

| | Activity Sets | | | | |
|---|---|---|---|---|---|
| **Work Coupling** | F-to-F | Proximal | Remote | Focused | Parallel |
| No Interaction | 0 | 3 | 1 | 77 | 44 |
| Light Weight | 1 | 28 | 13 | 4 | 32 |
| Information Share | 1 | 7 | 24 | 3 | 7 |
| Coordination | 10 | 18 | 38 | 15 | 9 |
| Collaboration | 82 | 42 | 23 | 4 | 7 |
| Cooperation | 7 | 2 | 1 | 0 | 1 |

The findings in Table 1 reflect a number of properties about activity awareness and its role in the functioning of distributed teams. First, students attempted more tightly coupled work when they interacted face to face and during proximal interaction. However, light-weight interaction was more prevalent during proximal activities than during face-to-face interaction. We speculate that the light-weight interaction we observed was due to the general familiarity of team members within a classroom. For example, students moved between light-weight interaction and other levels of work coupling much more fluidly during proximal interaction than during face-to-face work. Overall, there was considerably less face-to-face interaction (2%). The students were task oriented, and spent the majority of their time in the upper bands of work coupling in this activity. Spending more time in light-weight interaction during face-to-face activities likely would have increased if students had spent more time in this activity.

The work during remote interaction tended to be at the intermediate levels of coupling. Compared to F-to-F and proximal, there was more information sharing and coordination and less collaboration and cooperation. Although little interaction took place during focused work, it is interesting to note that students did use this time for coordination purposes. That is, we observed that students moved from personal work to interacting with others mostly for coordination purposes. When they did stop working individually, it was to situate their activities with the activities of others.

Students pursued a variety of goals during parallel activities, including interacting with teachers and students in the classroom who were not part of their group. While they were doing these things, they either had no interaction with their group members or interacted in a light-weight manner. Knowing people well and interacting face to face (proximal and parallel) promoted casual social interaction.

In year one, students had difficulty in completing their projects. Proximal members had few collaborative breakdowns, but remote interaction was plagued with awareness and common ground issues. Direct observation, interviewing, and questionnaires revealed that students struggled to understand what their remote partners were doing or why. Many difficulties were due to contextual factors particular to each classroom. Much of this information was shared during light-weight interactions. For example, students in one classroom failed to construct a physical model, collect data, and share it because some of the parts to the equipment kit were missing. The parts were due to arrive the following week, but the remote students never received this

information. The proximal members knew this was the case without having to share the information explicitly. The remote partners were unable to graph the data and stay on schedule according to the plan. This did more than disrupt planned activities for both groups. The remote partners felt the other members were not completing the work as agreed upon.

Distributed process loss for remote groups consisted of significantly more time in information sharing and coordination. Remote partners continually clarified what they were doing and how each respective side should proceed. This compromised their ability to work more closely together. Clearly, the face-to-face and proximal groups were able to spend more time in collaboration and cooperation. Note however that we should not conclude that co-present students experienced less of a demand to share information or coordinate activities. Rather, these students managed such needs as background tasks during other times. In this sense, the distributed process loss was less detrimental to the entire collaborative process.

By concentrating on the level of work coupling and the resulting communication, we were able to detect patterns that documented how demands on the communication and coordination process led to problems in common ground and awareness. Using collaborative technologies fractured contextual information critical to the collaborative process. Contextual information was needed but was poorly supported by the system. It was unusual for students to share background context unless collaborative breakdowns developed that indicated a need for it. Unfortunately, without adequate common ground and awareness, students often did not even understand the sources of their problems, and therefore, did not understand what information was needed. They often assumed they shared the same context with their remote partners. Students who had a great deal of direct contact had many more sources of information that led to greater levels of awareness and common ground. As a result, there were fewer demands on information sharing and coordination, and important time could be spent collaborating and cooperating. Below we offer several suggestions for carrying out the type of analyses reported here.

## 8. DISCUSSION
At the beginning of this paper, we outlined 3 difficult problems for evaluating distributed CSCW: (1) logistical difficulties in collecting data, (2) number and complexity of variables to consider, and (3) focusing on the re-engineering of work practices. These problems cannot be eliminated, but below we offer several suggestions for mitigating their effects.

The first step in any evaluation should be to clearly state the objectives and criteria of success by generating an evaluation plan. This information often does not get reported because standards for reporting findings are inconsistent across disciplines. Several properties must be defined: problem definition, purpose, type of evaluation, characteristics of data collection, focus of evaluation, conceptual framework, evaluation goals and objectives, research questions, and analysis. The research problem and goals and the resulting questions should drive the other factors. These two factors should then be mapped to methods using an *evaluation methodology matrix*. The matrix should be used to show which data collection and analysis techniques relate to which evaluation questions and goals. Doing these steps prior to data collection will alleviate issues associated with problems 2 and 3 listed above. The conceptual model plays a prominent role in insuring that the number of variables and their complexity are

addressed or not. Many texts offer very broad advantages and disadvantages of different methods [see 33], but it is almost non-existent to see reports generalizing which methods tackle which problems, especially in CSCW. Creating a literature base that maps methods to specific collaborative problems would be very useful to CSCW researchers. Collaborative problems could also be categorized as collaborative evaluation patterns that include methods useful for their study.

Each issue in a conceptual model should be studied with converging methods. We have continually advocated mixed-method designs [11]. Multiple techniques can be used to triangulate findings across components of a model. Multiple methods are needed to map the interactions at each stage in a model as well. It is the aggregate of findings that must be considered across each level in the model. A critical attribute in doing this is reconstructing distributed findings during the analysis stages. Only by reconstructing the circumstances of distributed settings can researchers understand the issues. To do this effectively, performance, process, and satisfaction measures must be used that have quantitative and qualitative properties. This is a *pragmatism* paradigm that is driven by the research problem, rather than by methods [44]. In a similar vein, we recommend this approach be used across laboratory and field settings. This can be done within or across researchers and projects.

We have been developing evaluation approaches that integrate laboratory studies and fieldwork [7, 18], and we believe the best approach is to explicitly and systematically combine the two. In particular, we recommend the combination of *simulation experiments* [24] and fieldwork in CSCW. Simulation experiments stage situations in the laboratory that are as similar as possible to the operational setting. The simulations are a compromise between highly controlled laboratory experiments and descriptive field studies. The simulations enable collection of precise and reliable data, while simulating realistic natural settings. Simulations enable controlled repeated observations, and the ability to study complex tasks over longer time periods than in formal experiments [51]. Of unique value is the ability to introduce disturbances or probes. This type of experimentation relies on well-developed scenarios that are derived directly from field data. Representativeness can be maximized when manipulating variables or system components to insure findings are generalizable to actual contexts of use [48].

The traditional view holds that dimensions identified in fieldwork are used to guide formalization, quantification, and experimentation in the lab; however, we subscribe to Xiao and Vicente's position that the process can be bidirectional, with top-down deduction and bottom-up abstraction informing each other [52]. The process of using mixed methods across laboratory and field studies significantly reduces all three of the problems identified above. Logistics of data collection can be dramatically reduced in the lab. The number of variables and their complexity can also be better managed. Lastly, specific research hypotheses can be identified in the field and targeted in the lab that focuses on the re-engineering of group work.

## 9. CONCLUSIONS
In this paper we have described several formidable challenges in evaluating computer-supported collaborative activities. In particular we have focused on complex, long-term activities carried out by teams of co-present and remote partners interacting

in a range of synchronous and asynchronous modes. Although there has been a great deal of work towards developing CSCW evaluation frameworks, our awareness evaluation model extends this work and provides a more refined model for remote collaboration in particular. Our model focuses on the processes of group work, namely communication, collaboration, and coordination. Further, we described how contextual factors, awareness, and common ground relate to and frame the processes of group work.

It is important that CSCW evaluators do not become a "human factors police." Both naturalistic and positivistic methods used in CSCW evaluation are seriously lacking in their appropriateness for producing design solutions. Few methods have been developed with creating engineering solutions in mind. It is possible, but researchers must be continually cognizant about how data collection and analysis methods will translate into design solutions.

We have argued that second-level social system effects based on properties of human coordination and collaboration are most likely to be relevant for CSCW evaluation. CSCW tools have profound effects on social systems, and this should be the basis for evaluating collaborative technologies. The deficiency in sound empirical methods for determining outcomes has been one of the leading causes for the lack of groupware success. Better evaluation approaches are critical to the successful development of CSCW applications. The work presented here is intended as an essential step in that direction.

## 10. ACKNOWLEDGMENTS

## 11. REFERENCES

[1] Arrow, H., McGrath, J. E., and Berdahl, J. L. (2000). *Small Groups as Complex Systems: Formation, Coordination, Development, and Adaptation*. Thousand Oaks, CA: Sage.

[2] Borghoff, U. M., and Schlichter, J. H. (2000). Computer-Supported Cooperative Work: Introduction to Distributed Applications: Springer.

[3] Brooks, F. P. (1995). The Mythical Man-Mouth. Reading, Massachusetts: Addison-Wesley.

[4] Carroll, J. M., Neale, D. C., Isenhour, P. L., Rosson, M. B., and McCrickard, D. S. (2003). Notification and awareness: synchronizing task-oriented collaborative activity. International Journal of Human-Computer Studies, 58, 605-632.

[5] Carroll, J. M., Chin, G., Rosson, M. B., & Neale, D. C. (2000). The development of cooperation: Five years of participatory design in the virtual school. In Proceedings on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (pp. 239-251), New York: Association for Computing Machinery

[6] Clark, H. H. (1996). Using Language. New York: Cambridge University Press.

[7] Convertino, G., Neale, D. C., Hobby, L., Carroll, J. M., and Rosson, M. B. (2004). A laboratory method for studying activity awareness. In NordiCHI 2004 (to appear). Tampere, Finland: Association of Computing Machinery.

[8] Damianos, L., Hirschman, L., Kozierok, R., Kurtz, J., Greenberg, A., Walls, K., Laskowski, S., and Scholtz, J. (1999). Evaluation for collaborative systems. ACM Computing Surveys, 31(2).

[9] Dourish, P., and Bellotti, V. (1992). Awareness and coordination in shared workspaces. In Proceedings of the ACM CSCW '92 Conference on Computer Supported Cooperative Work (pp. 107-113). New York: Association for Computing Machinery.

[10] Ganoe, C. H., Somervell, J. P., Neale, D. C., Isenhour, P. L., Carroll, J. M., Rosson, M. B., and McCrickard, D. S. (2003). Classroom BRIDGE: Using collaborative public and desktop timelines to support activity awareness. In Proceedings of the ACM '03 Symposium on User Interface Software and Technology (UIST) (pp. 21-30). New York: Association for Computing Machinery.

[11] Greene, J. C., Caracelli, V. J., and Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. Educational Evaluation and Policy Analysis, 11, 255-274.

[12] Grudin, J. (1988). Why groupware applications fail: Problems in design and evaluation. Office: Technology and People, 4(3), 245-264.

[13] Grudin, J. (2002). Group dynamics and ubiquitous computing. Communications of the ACM, 45(12), 74-78.

[14] Hare, A. P. (1992). Groups, teams, and social interaction: Theories and applications. New York: Praeger.

[15] Hartswood, M., and Procter, R. (2000). Design guidelines for dealing with breakdowns and repairs in collaborative work settings. International Journal of Human-Computer Studies, 53, 91-120.

[16] Helms, J., Neale, D. C., and Carroll, J. M. (2000). Data logging: Higher-level capturing and multi-level abstracting of user activities. In Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society (pp. 303-306). Santa Monica, CA: Human Factors and Ergonomics Society.

[17] Hendrick, H. W., and Kleiner, B. M. (2000). Macroergonomics: An introduction to work system design. (Vol. 2). Santa Monica, CA: Human Factors and Ergonomics Society.

[18] Humphries, W., Neale, D. C., McCrickard, D. S., and Carroll, J. M. (2004). Laboratory Simulation Methods for Studying Complex Collaborative Tasks. In Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society. Human Factors and Ergonomics Society (to appear). Santa Monica, CA.

[19] Isenhour, P. L., Carroll, J. M., Neale, D. C., Rosson, M. B., and Dunlap, D. R. (2000). The virtual school: An integrated collaborative environment for the classroom. Educational Technology and Society, Special Issue on "On-Line Collaborative Learning Environments", 3(3), http://ifets.ieee.org/periodical/.

[20] Klein, G. (2001). Features of team coordination. In M. McNeese, E. Salas, and M. Endsley (Eds.), New Trends in

Cooperative Activities: Understanding System Dynamics in Complex Environments (pp. 68-95). Santa Monica, CA: Human Factors and Ergonomics Society.

[21] Malone, T. W., and Crowston, K. (1994). The interdisciplinary study of coordination. ACM Computing Surveys, 26(1), 87-119.

[22] McGrath, J. E. (1984). Groups: Interaction and performance. Englewoods Cliffs, NJ: Prentice-Hall.

[23] McGrath, J. E. (1990). Time matters in groups. In J. Galegher, R. E. Kraut, and C. Egido (Eds.), Intellectual Teamwork: Social and Technological Foundations of Cooperative Work (pp. 23-61). Hillsdale, N. J.: Lawrence Erlbaum.

[24] McGrath, J. E. (1994). Methodology matters: Doing research in the behavioral and social sciences. In R. M. Baecker, J. Grudin, W. A. S. Buxton, and S. Greenberg (Eds.), Readings in human-computer interaction: Toward the year 2000 (pp. 152-169). San Francisco, CA: Morgan Kaufmann.

[25] Miles, M. B., and Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook. Thousand Oaks, CA: Sage.

[26] Nardi, B. A. E. (Ed.). (1996). Context and consciousness: Activity theory and human-computer interaction. Cambridge, MA: MIT Press.

[27] Neale, D. C., and Carroll, J. M. (1999). Multi-faceted evaluation for complex, distributed activities. In Proceedings of CSCL'99 Computer Supported Cooperative Learning (pp. 425-433). Mahwah, New Jersy: Lawrence Erlbaum.

[28] Neale, D. C., Dunlap, D. R., Isenhour, P., and Carroll, J. M. (2000). Collaborative critical incident development. In Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society (pp. 598-601). Santa Monica, CA: Human Factors and Ergonomics Society.

[30] Olson, G. M., and Olson, J. S. (1997). Research on computer supported cooperative work. In M. Helander, T. K. Landauer, and P. Prablu (Eds.), Handbook of Human-Computer Interaction (pp. 1433-1456). Amsterdam, The Netherlands: Elsevier Science.

[31] Olson, G. M., and Olson, J. S. (2000). Distance matters. Human-Computer Interaction, 15, 139-178.

[32] Olson, G. M., and Olson, J. S. (2001). Technology support for collaborative workgroups. In G. M. Olson, T. W. Malone, and J. B. Smith (Eds.), Coordination Theory and Collaboration Technology (pp. 559-584). Mahwah, N. J.: Lawrence Erlbaum.

[33] Pedhazur, E. J. (1991). Measurement, design, and analysis: An integrated approach. Hillsdale, New Jersey: Lawrence Erlbaum.

[34] Pinelle, D., and Gutwin, C. (2000). A review of groupware evaluations. In Proceedings of the 9th IEEE WETICE (pp. 86-91).

[35] Pinelle, D., Gutwin, C., and Greenberg, A. (2003). Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the mechanics of collaboration. Transactions on Computer-Human Interaction, 10(4), 281-311.

[36] Pinsonneault, A., and Kraemer, K. L. (1989). The impact of technological support on groups: An assessment of the empirical research. Decision Support Systems, 5(2), 197-211.

[37] Plowman, L., Rogers, Y., and Ramage, M. (1995). What are workplace studies for?. In Proceedings of ECSCW '95 European Conference on Computer-Supported Cooperative Work (pp. 309-324). Dordrecht: Kluwer.

[38] Ross, S., Ramage, M., and Rogers, Y. (1995). PETRA: Participatory evaluation through redesign and analysis. Interacting with Computers, 7(4), 335-360.

[39] Ruhleder, K., and Jordan, B. (1998). Video-based interaction analysis (VBIA) in distributed settings: A tool for analyzing multiple-site, technology-supported interactions. In Proceeding of the Participatory Design Conference (pp. 195-196). Seattle, WA.

[40] Salvador, T., Scholtz, J., and Larson, J. (1996). The denver model for groupware design. SIGCHI Bulletin, 28(1), 52-58.

[41] Sproull, L., and Kiesler, S. (1991). Connections: New ways of working in the networked organization. London: MIT Press.

[42] Stanney, K. M., and Maxey, J. (1997). Socially centered design. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics (2nd Edition) (pp. 637-656). New York: Wiley.

[43] Suchman, L. (1987). Plans and situated actions. New York: Cambridge University Press.

[44] Tashakkori, A., and Teddlie, C. (1998). Mixed methodology: Combining qualitative and quantitative approaches. Thousand Oaks, CA: SAGE.

[45] Thomas, P. J. (Ed.). (1996). CSCW Requirements and Evaluation. London: Springer.

[46] Twidale, M., Randall, D., and Bentley, R. (1994). Situated evaluation for cooperative systems. In Proceedings of the ACM CSCW '94 Conference on Computer Supported Cooperative Work (pp. 441-452). New York: Association for Computing Machinery.

[47] Vicente, K. J. (1997). Heeding the legacy of Meister, Brunswik, & Gibson: Toward a broader view of human factors research. Human Factors, 39(2), 323-328.

[48] Watts, L., and Monk, A. (1996). Inter-personal awareness and synchronization: Assessing the value of communication technologies. International Journal of Human-Computer Studies, 44, 849-873.

[49] Whittaker, S., and Schwarz, H. (1999). Meetings of the board: The impact of scheduling medium on long term group coordination in software development. Computer Supported Cooperative Work, 8, 175-205.

[50] Woods, D. D. (2003). Discovering how distributed cognitive systems work. In E. Hollnagel (Ed.), Handbook of cognitive task design (pp. 37-53). Mahwah, N.J.: Lawarence Erlbaum.

[51] Xiao, Y., and Vicente, K. J. (2000). A framework for epistemological analysis in empirical (laboratory and field) studies. Human Factors, 42(1), 87-101.