In *Proceedings, Requirements Engineering '93*, edited by Stephen Fickas and Anthony Finkelstein, IEEE Computer Society, 1993, pages 152–164.

Techniques for Requirements Elicitation[1]

Joseph A. Goguen
Centre for Requirements and Foundations
Oxford University Computing Lab

Charlotte Linde
Institute for Research on Learning
Palo Alto, California

Abstract

This paper surveys and evaluates some techniques for eliciting requirements of computer-based systems, paying particular attention to how they deal with social issues. The methods surveyed include introspection, interviews, questionnaires, and protocol, conversation, interaction, and discourse analyses. Although they are relatively untried in Requirements Engineering, we believe there is much promise in the last three techniques, which grew out of ethnomethodology and sociolinguistics. In particular, they can elicit tacit knowledge by observing actual interactions in the workplace, and can also be applied to the system development process itself.

1 Introduction

A basic question in Requirements Engineering is how to find out what users really need. Research has shown that many large projects fail because of inadequate requirements [5]; moreover, this inadequacy is often related to social, political and cultural factors. This paper describes and assesses techniques for requirements elicitation. We first review some traditional techniques, including introspection, questionnaires, interviews, focus groups, and protocol analysis. Then we discuss some techniques from discourse analysis, which we take to include conversation and interaction analyses, as well as the analysis of discourse structure. Finally, we compare the various approaches. There is a fairly large bibliography.

Developing a large system[2] is a complex and difficult process. In the early days of computing, there was no particular organisation to this process: programmers just sat down and tried to write code that would be useful. Today, few doubt that a task that can consume hundreds of person-years should be carefully planned and managed. Therefore the system "life cycle" has been broken into a number of so called "phases," of which Requirements Engineering is the earliest phase[3] that lies largely within Computing Science. The requirements phase is typically preceeded by business planning, and is formally initiated by the client.

It is more accurate to view the division of the life cycle into phases as a management technique, than as a model of how the system development process actually proceeds; i.e., the life cycle phases are a useful scheme for classifying the activities that occur in system development, but it is far from true that these activities occur in strict linear order. Suchman [43] explains that naturally occurring plans are typically used as after-the-fact explanations to lend coherence to past events. Indeed, requirements are constantly reconsidered in both design and coding, and often activities that can be classified as Requirements Engineering are done by programmers and managers relatively near system delivery, or even after system delivery. Moreover, much of this work remains undocumented. (See [8] for an ethnographic study that supports these assertions.)

Once a need is expressed and an initial plan developed, the requirements team tries to identify what properties the system should have to meet that need. Note that setting up a requirements team involves choosing representatives of the client; their background knowledge and experience can play a very strong rôle in the development process. Relevant properties may include not just high-level functional requirements, but also response time, cost, security, portability, reliability, and modifiability. In addition, there may be requirements for the development process, such as certain quality control procedures, reporting schemes, tools, or limits on cost or time. Some of these are not easily quantified; however, some imprecision may even be desirable, to accommodate the trade-offs that inevitably arise.

The next phase after requirements is "design," where engineers try to fix the main components of the system, their requirements, and interactions. This resembles how an architect designs a house, once requirements have been agreed with a client. A more detailed design phase may follow. (Of course, an actual execution of this idealised plan will generally interleave the various activities.)

The analogy with architecture should make it clear that eliciting requirements can be far from easy: clients may change their minds once they see the possibilities more clearly, and discoveries made during later phases may also force retrofitting requirements. The requirements of real systems are rarely static. There are very good reasons why clients often do not, or cannot, know exactly what they need; they may want to see models, explore alternatives, and envision new possibilities. Often

[2] There is little difference between the development cycles of software and hardware systems, and most large real systems involve both aspects.

[3] There is no widely accepted terminology for phases, nor even any widely accepted division into phases.

these possibilities are closely intertwined with social, political, legal, financial, and/or psychological factors. For example, certain ways of using a database may be illegal; others may be politically undesirable; some may be incompatible with the corporate organisation of the user (e.g., they may cross administrative boundaries); others may be too slow unless very expensive equipment is used. In the extreme, a project may be doomed, because no system can be built that satisfies its requirements or because the agreed requirements do not reflect the real needs.

A major goal of Requirements Engineering is to avoid such problems. This will often involve putting significant effort into requirements elicitation. Unfortunately, Requirements Engineering is an immature discipline, perhaps not entirely unfairly characterised as a battlefield occupied by competing commercial methods[4], firing competing claims at each other, and leaving the consumers weary and confused.

1.1 Why Social Science?

The problems of requirements elicitation cannot be solved in a purely technological way, because social context is much more crucial than in the programming, specification and design phases. Some Computing Scientists might think that requirements elicitation is where science stops and chaos begins. This raises the fundamental questions of whether there *is* any order in the social world, and if so, how it can be studied. If there is order in the social world, then a precise understanding of how it is constructed and maintained should help with methodology for requirements elicitation. If not, then requirements elicitation must remain a mysterious process, fraught with frequent unexplained failures, and occasional unexplained successes.

The premise of this paper, as of social science generally, is that the social world *is* ordered. We also make two further assertions: social order may not be immediately obvious, or immediately describable, by common sense; and social order cannot be assumed to have an *a priori* structure. Therefore, social order can only be determined by immersion in the actual unfolding of social phenomena, rather than (for example) by collecting statistics about the occurrence of certain pre-given categories. Detailed arguments for these assertions are given later.

The majority of computer-based systems are developed without any systematic help from the social sciences (sociology, psychology, linguistics, anthropology, etc.). This means that the needs of the user, both as individual and as organisation, are not addressed systematically; in general, they are only incompletely known to the development team, and there are often some serious misconceptions. Among the systems that have been developed with some help from the social sciences, most have used only classical experimental psychology (e.g., ergonomics for keyboard layout, or the psychology of perception for display colours). Many efforts have tried to model the cognitive process of individual users, but this approach has not been very successful with the larger social, political, and cultural factors that so often cause failure.

Very few system development efforts have tried to use any social science methods beyond (for example) elementary guidelines for the conduct of interviews. Among these, very few indeed have tried to use techniques based on what we regard as the most promising areas, namely ethnomethodology and sociolinguistics (see [17, 28] for related discussion).

It seems worth emphasising that many requirements methods available in the marketplace, even though they may refer to certain social, organisation, or linguistic issues, do not do so in a systematic manner, and in fact, do not have any proper scientific basis at all. Of course, this is not to deny that there may be a great deal of practical experience behind the recommendations and notations of some of the better methods, or that they may be useful in many practical situations.

This paper begins to explore a scientific basis for requirements elicitation, by considering the basic issue of how to acquire the necessary information. Introspection is undoubtedly the most common current source of information; but experience shows that it can be very misleading. Interviews and questionnaires are also widely used, and sometimes protocol analysis is used. Any of these can be useful. But this paper argues that conversation, interaction, and discourse analyses are more detailed and precise, and hence likely to be more accurate.

Acknowledgements

The first author wishes to thank Kathleen Goguen for many valuable comments on this paper, and the members of the Centre for Requirements and Foundations at Oxford for their friendly enthusiasm. Both authors wish to thank Dr. Susan Leigh Star for her very helpful comments on a draft of the paper.

2 Introspection

Introspection is the first and most obvious method for trying to understand what properties a system should have in order to succeed. It amounts to imagining what kind of system I would want if I were doing this job, using this equipment, etc. This method can be very useful, but it has the problem that the introspection of an expert in a different field, such as Requirements Engineering, is unlikely to reflect the experience of actual users. Experts tend to work from what they remember or imagine of themselves; for user interface design, this experience can

[4] What we call "methods" are often called "methodologies" by practitioners. But in an academic context, the word "methodology" should properly be used for the study and comparison of methods, and that is how we use it in this paper.

be very far from the questions, assumptions and fears of actual users. In viewing tapes of novice users learning a new interface, interface designers and cognitive scientists were consistently shocked by what they saw as incompetent and inconsistent behaviour [25]. For example, experts might be surprised that, when a word processer does something a user finds surprising, such as centering a headline further to the right than expected, the user does not attempt to understand why; in fact, users seem to believe that computers just are sometimes puzzling or irritating, and that it is not necessary or valuable to explain why. Cognitive scientists may be surprised at this, because their model suggests that a user who finds that a model is incorrect should correct the model. Designers may be upset because they feel that the subjects are not using their designs correctly.

Similarly, requirements engineers cannot introspect what work settings look like, or the conditions under which a new technology will be learned. For example, many subjects must learn to use new technology in conditions that require multiple and ongoing splitting of attention. However, requirements rarely take account of this.

Finally, we note that the phrase "naive user" can confuse the issue. So called naive users are often experts in their own speciality, about which the requirements engineers are naive. Although this is obvious, the point is that the phrase focuses attention on the users' relation to the new technology, and may suggest that the users' task is to learn the technology properly and fully, instead of just doing their own job better using the new technology.

We conclude not that introspection is an inadmissible method, as claimed by many current psychologists, but rather that introspection without careful consideration of its limits can be (and often is) highly inaccurate (an interesting discussion of introspection can be found in [47]). Hence, we suggest that if there is room for doubt, introspection should be checked by some of the more empirical methods described below.

3 Interviews

Interviews are used in an extraordinary variety of domains, and are often quite successful; see [30] for a good survey. This section discusses questionnaires, open ended interviews, and focus groups, showing that the interview process involves some (usually unstated) assumptions about the interaction between interviewer and subject. We argue that some of these assumptions are quite problematic, and raise doubts about using these methods for some applications.

3.1 Questionnaire Interviews

Questionnaire interviews are very widely used, and have the benefit of appearing scientific, because they use statistical analysis. The following is from a discussion by Suchman and Jordan ([46] p. 232):

1. There is an unresolved tension between the survey interview as an interactional event and as a neutral measurement instrument. On the one hand, the interview is commonly acknowledged to be fundamentally an interaction. On the other hand, in the interest of turning the interview into an instrument, many of the interactional resources of ordinary conversation are disallowed.

2. The success of the interview as an instrument turns on the premise that (a) relevant questions can be decided in advance of the interaction and (b) questions can be phrased in such a way that, as long as they are read without variation, they will be heard in the intended way and will stimulate a valid response.

3. The premises of 2. fail insofar as (a) topics that come from outside the conversation run the risk of irrelevance, and (b) as an ordinary language procedure, the survey interview is inherently available for multiple interpretations of the meaning of both questions and answers.

4. Compared with ordinary conversation, the survey interview suppresses those interactional resources that routinely mediate uncertainties of relevance and interpretation.

Suchman and Jordan [46] argue that validity is not assured by having the same words repeated to subjects in each interview, because these words will mean different things to different people in different contexts. In normal interaction, these issues of interpretation are negotiated between participants; but in a survey interview, the method and training given to interviewers specifically forbids such negotiation. The following example should make this point more vivid [46][5]:

I: Generally speaking, do you usually think of yourself as a Republican, Democrat, Independent, or what?

R: As a person.
I: As a Republican::
R: No.
I: Democrat::
R: No.
I: Independent or what.
R: Uhm:: I think of myself as a (pause) Christian.
I: Okay. (writing) But politically, would you have any particular:: (inaudible)

[5] In this transcription system, colons after a sound indicate that it is lengthened, and the number of colons indicates the degree of lengthening.

R: I am one of Jehovah's Witnesses so, you know, when it comes to::

I: I see.

R: So I'm, I am acclimated toward government, but it is that of Jehovah God's kingdom.

I: Yes.

Here, the interviewer presupposes a system of political categories, and asks the respondent to choose one for self identity. But the respondent does not share this system, and thus cannot choose. This mismatch could be the beginning of an interesting exploration of the respondent's religious and political categories, and in an ordinary conversational situation, probably would be. But because this is forbidden for survey interviewers, this fascinating informant probably ends up as a bleached "Don't Know" or "Other." The point is a general one: categories and concepts that are transparent to one community can be entirely opaque to members of another community, and the fact that this opacity exists may not be noticed in the course of discussions unless specific attention is paid to the possibility.

Here is another example (cited in [30]) of an answer that must be classified as "other":

I: Are you a virgin?

R: Not yet.

3.2 Open Ended Interviews

The open ended interview is much used in anthropology and psychology, and avoids many problems of the questionnaire method. In it, the interviewer poses a question, and then allows subjects to answer as they wish. The interviewer may probe for more detail, but does not set the terms of the interview. This sounds much more benign than the survey interview, but the issues of whether the question asked can be answered at all, and whether the answer is part of the normal discourse repertoire of the speaker, still remain.

Let us first consider questions that cannot be answered at all. For example, in linguistics and education research, subjects are sometimes asked how they tie their shoelaces. This produces some marvelous examples of linguistic incompetence. But there is no reason why subjects *should* be competent at this task, because people do not tell each other how to tie bowknots — rather, it is taught by showing. (But a sailor or a ship model maker may give a much more competent performance, because these experts have vocabulary for knots and the parts of knots.)

Let us generalise. People know how to do many things that they cannot describe. It is a commonplace in ethnography that people's descriptions of how they weave a basket or choose a chief or write a program bear a complex and opaque relation to how they can be seen to do these things when they are observed. This problem is so familiar that it has a nickname in social science: the *say-do problem*; also, philosophers speak of *tacit knowledge*. The moral is this: Don't ask people to describe activities that they do not normally describe, or if you do, then don't believe the answers.

Now let us consider interview questions that people can and do answer in what seems a useful way, and ask how this compares with their practice; we must consider not only the practice that they describe, but also their discourse practice. For this, we must compare the discourse produced when the topic is elicited with that produced in a related but non-elicited situation; that is, we ask whether the interview data is the same, is wholly different, or bears some partial but regular similarity to the non-elicited speech. One approach is to observe spontaneous speech. For example, in studying apartment layout descriptions, we can observe whether they occur in spontaneous speech, and whether they are the same or closely similar to instances gathered in an interview situation [23, 27]; such informal checking must be done after the actual analysis of elicited data has suggested some structures of interest, because memory for linguistic structure in natural settings is generally not sufficiently reliable. In practice, such observation requires interest in the adventures of friends searching for apartments, and following strangers down the street when their conversation turns to this topic. Such are the exigencies of empirical research.

It is even better to compare interview data with recorded, non-elicited data, to see if they are usably similar. For example, [51] shows that elicited narratives differ from spontaneously produced narratives on a fine level of detail, including use of the historical present tense. This difference arises because *performing* a narrative, so that the addressee can visualise the event, encourages use of the historical present. Performed narratives are much more likely to be produced when the speaker and addressee share characteristics such as age, occupation or ethnicity, or when they are friends. Because these characteristics are not likely to be shared by participants in an interview situation, the tense system will be at least slightly different. Similarly, the form of evaluation (see Section 6.2) in elicited narratives may differ from that of spontaneous narratives, because spontaneous narratives can include negotiations between the primary speaker and other interlocutors that an interviewer may be unwilling to undertake, for fear of biasing the data.

Whether such differences matter depends on the nature of the investigation, and must be determined for each case. If finely detailed data is needed, then elicited narratives cannot be considered identical to spontaneous narratives. But if only less detailed, or higher level structure is needed, then open ended interview data may be adequate.

3.3 Focus and Application Development Groups

The focus group is a kind of group interview, rather widely used in marketing research [42], but less used in pure social science research ([29] gives a favourable review of the potential of focus groups in social science). In this technique, groups are brought together to discuss some topic of interest to the researcher. In market research, this is often done using stimulus materials such as films, story boards, or product mockups as a focus (hence the name), and is commonly used to get the opinions of representative potential customers on new products.

Focus groups have the advantage of allowing more natural interactions between people than questionnaire interviews, or even open ended interviews. However, the groups are usually not natural communities, such as people who eat lunch together, or all the purchasing agents of a particular corporation, but rather are an *ad hoc* collection, constituted for the occasion by the researcher, usually on the basis of demographic considerations. Further, although focus groups may be valuable for eliciting responses to products whose features and trade-offs customers understand (for example, whether they would be willing to pay more for upscale gourmet dog food for their Dobermann Pinschers), they are not useful in eliciting opinions on design issues where the subjects are not experts, and therefore must respond within the categories and structures provided by the researcher.

So called JAD or RAD groups[6] have recently become popular in Requirements Engineering, especially for Information Systems applications, because of their claim to greatly accelerate the development of requirements [1]. This method is closely related to focus groups, and can be expected to suffer from some of the same problems. In particular, participants will certainly be unable to articulate tacit knowledge. Also, even though group facilitators try to avoid imposing their own categories on participants, there is no guarantee that the participants will in fact share categories with each other. Moreover, because participants may have widely different status within the organisation, there is a danger that some will not feel free to say what they really think, especially if it is unpopular. Finally, it will often be difficult for non-technical participants to assess the significance of technical decisions. Although this method appears promising, we believe its potential limitations should be studied empirically.

3.4 Discussion

Interview methods can fail if the interviewer and respondent do not share a category system. For example, because the clients of architects are usually unfamiliar with the conventions of architectural drawings, they can easily agree to a design that fails to satisfy them when built. The Workplace Project[7] [7, 44, 45] found a case where some workers who wanted a space for working together, agreed to a protruding addition to a long tabletop. However, when built, the angle of the protrusion made it harder to view documents jointly than before.

Similar issues arise in requirements elicitation, because requirements engineers often come from communities with different values, assumptions, concerns, etc. from those of users. For example, [6] describes a case where two management information consultants working for a large U.S. university encounter difficulties in promoting schemes to help students, and then explain those difficulties with the theory that the administration is really concerned with "the care and feeding of the faculty," but cannot say so because this conflicts with the university's official mission statement.

Questionnaires, administered either orally or in writing, are often used in Requirements Engineering to determine characteristics or concerns of user populations. They can be useful when the population is large enough, and the issues addressed are clear enough to all concerned. However, they will fail when subjects are asked about topics that they do not have ways to talk about, or do not want to talk about.

4 Protocol Analysis

Protocol analysis asks a subject to engage in some task and concurrently talk aloud, explaining his/her thought process. Proponents claim that this kind of language can be considered a "direct verbalization of specific cognitive processes" ([12], p. 16). Protocol analysis is also used to reflect on problem solving, or some other task, retrospectively, i.e., after it has been accomplished. This section considers concurrent talk-aloud protocols, because they are more common; however, we note that the arguments of previous sections apply to the retrospective approach.

There seem to be two main arguments for talk-aloud protocols: that they are possible, and that they work. The argument for possibility must overcome arguments in psychology about the method of introspection used in the nineteenth and early twentieth centuries. It is claimed [32] that the arguments against introspection do not apply, because the subject is not introspecting, but rather is emitting a stream of behaviour that does not differ in kind from producing a galvanic skin response or a muscular movement. Therefore, talk-aloud protocols are possible data in the framework of behaviourist psychology.

The argument that protocol analysis works in [12] is based on the apparent success of the GPS (General Problem Solver) system. GPS was originally developed as a research vehicle for problem solving in Artificial Intelli-

[6] These acronyms stand for Joint Application Development and Rapid Application Development, respectively.

[7] This study, conducted by Xerox Palo Alto Research Center, used as its data an airline operations room that was redesigned for a move to a new terminal.

gence. It reduces goals to subgoals, and then attempts to solve each subgoal that remains open by reducing it to further subgoals, until all subgoals are solved. In [32], the goals are to deduce certain symbolic logic expressions from others, and the transformations are certain elementary steps of deduction[8].

The results of this research have not stood the test of time. Even within the Artificial Intelligence community, it is hard to find researchers who still believe in top down backtrack problem solving that is driven simply by matching rules to goals to generate subgoals. For example, in motion planning, e.g., for robots, it is becoming increasingly clear that developing a complete plan in advance of execution is difficult, inefficient, inflexible, and does not correspond to how humans carry out actions. For the first point, studies in computational complexity show that producing complete, precise mathematical plans is an intractable problem [9]. For the second, it has been found that inexact, heuristic methods work better; indeed, so called "opportunistic planning," which produces partial plans, and then incrementally replans in response to new information, works much better in practice, because the sensory information and background knowledge available to robots (as to humans) is generally inexact, incomplete, ongoing, and subject to change. For a detailed summary of recent research on human planning, see [43].

Even in mechanical theorem proving, where the difficulties of being embedded in the physical world do not arise, current research uses a variety of heuristics, and also employs techniques that support flexible replanning — as do human mathematicians. No current generation Artificial Intelligence systems bear more than a superficial resemblance to GPS. Prolog [10], which at first glance might seem similar, in fact differs greatly in that it has logical variables, unification, cut, and other non-logical features to force evaluations that are *not* top down. (See [4] for a survey of mechanical theorem proving systems.)

In fact, protocols were *not* used inductively for developing GPS; rather, GPS was developed on the basis of *a priori* principles about mathematical problem solving, and then used as a basis for describing and critiquing the arguments in empirically obtained protocols. Of course, GPS was a significant advance in its time, and indeed, had to be explored thoroughly before moving on. We have no wish to minimise its historical importance in Artificial Intelligence, or its influence on cognitive psychology, both of which were considerable. However, we *do* wish to point out that its claims about human problem solving were wrong, even allowing a narrow interpretation of its domain, as were its claims about efficient mechanical problem solving. This implies that the experimental method used must have been flawed, and that is our main point: protocol analysis is not a reliable guide to what subjects are thinking, and is open to serious misinterpretation by analysts, who can choose a small sample of protocols (just one was used in [32]!) for an unrealistic problem (both artificially simple and artificially without social context) to impose their preconceptions on the data.

Beyond this demonstration of its fallibility, one can give two further arguments against protocol analysis, one general, and one specific to requirements. The first argument is this: As we have said, the assumption in staging and studying protocols is that people can produce language that gives a trace of autonomous cognitive activity. The problem with this assumption is that language is intrinsically social, created for a partner in conversation. (This property is called *recipient design* in conversation analysis.) When an experimenter asks a person to solve a problem and talk aloud, then that person has to imagine an experimenter with certain desires, and try to provide what the experimenter wants. (Or the subject may be rebellious, and try to frustrate the imagined experimenter.) Thus, protocols are an unnatural discourse form, and moreover, are unnatural in ways that are difficult to specify.

Let us consider the protocol given by Newell and Simon [32], produced by a student doing a problem in elementary symbolic logic:

> Well, looking at the left hand side of the equation, first we want to eliminate one of the sides by using rule 8. It appears to be too complicated to work with first. Now — no, — no I can't do that because I will be eliminating either the Q or the P in that total expression. I won't do that at first. Now I'm looking for a way to get rid of the horseshoe inside the two brackets that appear on the left and right sides of the equation. And I don't see it. Yeh, if you apply rule 6 to both sides of the equation, from there I'm going to see if I can apply rule 7.
>
> I can almost apply rule 7, but one R needs a tilde. So I'll have to look for another rule. I'm going to see if I can change that R to a tilde R. As a matter of fact, I should have used rule 6 on only the left hand side of the equation. So use rule 6, but only on the left hand side.
>
> Now I'll apply rule 7 as it is expressed. Both — excuse me, excuse me, it can't be done because of the horseshoe. So — now I'm looking — scanning the rules here for a second, and seeing if I can change the R to a tilde R in the second equation, but don't see any way of doing it. (Sigh.) I'm just sort of lost for a second.

There are specific linguistic features demonstrating the unnatural provenance of this passage. First, it fluctuates between the language of talking to oneself and the language of talking to an interlocutor who is physically present and involved. One sign of this fluctuation is the shift in pronoun choice: "we," "I," and "you" all appear. "You" seems to be the first choice in language that is specifically produced to be understood by an overhearer

[8] The anthropomorphic language used here is just a convenient shorthand for sketching the design of a computer program.

as talking to oneself. A first person plural expression like "first we want to eliminate one of the sides by using rule 8" is much more characteristic of a lecturer talking to an audience. Similarly, the use of impersonal constructions, such as "Well, looking at the left hand side" in the first sentence, and "It appears to be" in the second, is characteristic of the language of successful science [20]. Phrases like "excuse me, excuse me, it can't be done because of the horseshoe" are produced for an interlocutor; it is incoherent for a speaker to provide this kind of excuse to him/herself. Finally, a phrase like "I'm just sort of lost for a second" may most naturally be interpreted as an excuse to an interlocutor for a pause. In particular, note that the phrase "for a second" functions as a mitigation of the difficulty, as a kind of excuse for the pause, rather than as a bare description of a mental state.

The most telling argument against protocol analysis is that it *does not work*, as demonstrated by its use to support GPS as a model of human problem solving, once considered a spectacular success, but now seen to be a failure. Moreover, protocol analysis is based on a simplistic cognitivist model of human thinking as essentially computational, involving abstract representations of concepts, and their transformation by algorithms that are precisely specified by computer programs (e.g., see [31]).

Finally, even if it were possible to get a trace of a speaker's autonomous cognitive activity, such an object would be inappropriate for the requirements process, because the client does not have any pre-existing mental model of the desired system. Rather, the client has knowledge about business and organisational needs, while the requirements team has knowledge about technical possibilities. The process of producing requirements from these two different kinds of knowledge is necessarily conversational, because they must be combined. Thus, the requirements problem is intrinsically social, and cannot be solved using only methods that take individual cognition as fundamental.

5 The Question of Social Order

We have now surveyed a number of methods, and discussed some problems that arise from their underlying assumptions. The methods surveyed so far all impose an analyst's order on the social world, with no guarantee that this is the same as the order that members perceive, and with no way of even posing this as a research question. Note that the question of whose social order is assumed can be significant in requirements elicitation, where people from two or more possibly very different communities try to craft an understanding that is workable for all of them. An interesting discussion of some communication difficulties between sociologists and computer scientists is given in [36], based on actual experience at Lancaster University; one source of these difficulties appears to be the very different assumptions made in these two communities about the nature of research.

We have previously asserted the orderliness of the social world as a working principle; we will now examine in detail how certain aspects of social order are produced. Conventional approaches in sociology, anthropology, and the other social sciences assume pre-existing categories, such as social class, norm, rôle, etc., and then explain the observed social order as a reflection of these categories in practice. However, this approach does not explain how pre-given categories can act upon the moment to moment world of practice, to produce the order we observe. This critique of traditional social science is relevant to requirements elicitation because most existing approaches are based on it. If we are right, then the results of requirements studies that assume pre-given categories can easily be more inaccurate and misleading than necessary.

Ethnomethodology [13] and conversation analysis (which grew out of ethnomethodology) arose in response to these problems. These fields consider that social order is accomplished by members in their moment by moment activities. For example, consider a seminar. Although the word "seminar" suggests a pre-existing category, it is in fact constructed by the members' furnishing of a room, or choice of a room furnished in a certain way, in the arrangement of chairs, in the orientation of participants towards someone understood to be the speaker, in the allotment of a very long turn to the speaker, etc. It is the work of the participants that makes a seminar, not the category of "seminar" that makes the participants behave in a specified way. The view that social order is constructed by participants' actions, rather than being a pre-existing category that shapes people's actions, may be unfamiliar to many readers, and adopting it may require a different approach to studying social phenomena. This section discusses some fundamental premises that underlie ethnomethodology, using examples from a variety of fields, because the necessary research has not yet been done in the Requirements Engineering setting.

5.1 Natural Setting

To understand social order as an accomplishment of participants, we must study it in natural settings. A laboratory setting is constructed by an experimenter for a particular purpose, and it is considered bad experimental technique to reveal that purpose to the subjects. However, because humans are above all sense-making animals, they do not just sit with blank minds in a white room, passively enduring whatever comes. Rather, they continuously try to construct an understanding of the situation they are in, and then use this understanding to shape their behaviour as participants in the experiment, whether cooperative or subversive. Although the experimenter has control over the experimental setting, this does not determine what kind of sense their subjects make of it. Therefore, we may not get reliable results on the situation the experiment was intended to elucidate, because

we do not know what setting the participants think they are in, and their construction may well be very different from the setting that the experimenter had in mind.

For example, early studies of American black children's language argued that these children had a language deficit, and that sound educational policy required teaching them how to speak [3]. This research was based on evidence from experimental settings, in which a single child was brought into a room with an adult experimenter, usually white, shown some toy, like a plastic spaceship, and asked, "Now Johnny, I want you to tell me everything you can about this spaceship." In this context, the children tended to give short, simple, minimal descriptions, with an uncertain intonation, such as [3] "It's red? [Long pause] An' uh [pause] it's pointy?"

Looking at such responses, especially in contrast with the fluent responses of middle class white children, one might well be tempted to say that the black children needed to be taught how to talk. However, a different view was taken by Labov [21]. He went into a classroom with a rabbit and tape recorder, and told the children that the rabbit was shy and needed to be talked to so that it wouldn't be frightened. Then he and the teacher left the room. The language produced in this setting was extensive, fluent, and of startlingly greater complexity and competence than in the artificial test situation described above.

Such findings raise two important questions: "What is the difference between the two settings?" and "What causes the difference between the performance of black and white children in the artificial test setting?" The original test setting is so familiar and unproblematic to academics, who have had a lifetime of dealing with it, that we must pause to consider what it might mean to a black child. The child is asked to describe an object to a questioner who is at least as capable of seeing and describing it, because he owns the object. This is very different from the most common form of question, where the speaker does not know the answer, and has reason to believe that the hearer may. In a situation of such an enormous power differential — black child and white adult — the child in fact shows considerable social understanding in deciding that minimal talk is the least dangerous policy. That is, because the child does not understand the desire of the experimenter, he cannot construct the appropriate response, which in this case is to describe the object as if the experimenter could not see it and had never seen it before. It might be objected that white middle class children of a similar age can do just this kind of task. However, this does not show that their language abilities in general are greater. Rather, there is evidence that middle class white parents train children in just such decontextualised descriptions as a preparation for school: "Look at the kitty. What colour is the kitty?" Because the mother can see the kitty as well as the child, she does not need to be told that it is grey. But she is preparing the child for this kind of school question, which is decontextualised from the relation between speaker and hearer, their relative states of knowledge, etc. ([40], pp. 57–98).

5.2 Member's Categories

Perhaps the most important notion underlying the analysis of social order is that of *member's categories*. This notion comes from ethnomethodology and conversation analysis. The idea is to find the categories that members themselves use to order their social world, rather than to impose an analyst's order on it. For example, it is not useful to approach a given piece of interaction with the assumption that participants are doing a shockingly bad job of whatever it is the analyst decides they are doing. Rather, it is important to determine what the participants are actually doing. The fundamental idea is that the social world is already orderly, and this order is an on-going creation of the participants. Further, we as analysts don't know in advance what the relevant categories are, so we should not come to the data with a pre-given coding scheme.

For example [25], consider a party of eight people at a restaurant after a conference session. An analyst could use any number of category systems, e.g., three Xerox employees and five non-Xerox employees, seven employees of large organisations and one self-employed person, four people who had just given a paper and four who had not, one person who was pregnant, and seven who were not, four with blue eyes and four with brown eyes, six people who drank and two who did not, or one man and seven women. And of course this list could be multiplied indefinitely. The analyst needs to know what categories are relevant, and what relevance might mean. The notion of members' categories implies that the analyst should consider what categories the members themselves use to organise their interaction, that is, what categories they *orient to*. Thus in this situation, participants oriented to the category of pregnancy or non-pregnancy in deciding whether to take a taxi to the restaurant. They oriented to the nature of the participants' employers in determining what kinds of receipts were required. The waiter oriented to gender (and to recent developments in understanding the economic consequences of gender) in placing the bill in the center of the table facing the one man in the party, but not within his immediate reach. There was no evidence that eye colour was an organising category for any activity.

Let us consider further what it means for a phenomenon to organise an activity. The analyst should state what level of activity is of concern. There are striking phenomena that do not organise interaction at any level we care about. For example, a video of people interacting may clearly reveal particular ways that women with long nails and manicures use their hands, to protect their nail polish, which is easily chipped. But there is no evidence that other participants relate to this way of using the hands, or orient to it in organising their interaction. For

example, they do not pass objects to one another differently depending on whether the recipient has a manicure. However, an analysis of this way of using the hands may be very relevant to the design of certain products, particularly packaging.

It is implicit in the notion of members' categories as organising activity that analysts do not reconstruct intentions or mental processes, except in so far as these are evident to those involved in the activity. Thus, if someone starts writing on the upper left corner of a white board, we can say that this action projects that the board will probably be covered.

Another example is body torque: a posture in which and legs face front while the head and shoulders, or head shoulders and trunk, are turned sideways. This posture requires considerable muscular tension to hold for a long period. Therefore, conversations in torque, in which one interlocutor is partially turned towards another, are likely to be short. Thus, if a visitor walks into the office of someone working at a terminal, and the occupant turns his head and neck to greet the visitor while leaving his hands on the keyboard, the visitor can project that a short conversation is likely [18, 39].

This illustrates the demonstration of intention that is needed for this kind of analysis. Analysts cannot simply construct subjects' mental models or intentions. Rather, it is necessary to demonstrate what participants are doing that allows other participants to infer their intentions. Thus, the activity of the analyst in postulating intentions is not different from that of the participants, and proceeds on the same evidence. This leads to the discussion of members' methods.

5.3 Member's Methods

Suppose you are a musician who wishes to study Balinese music. One approach is to transcribe Balinese pieces on Western music paper, based on the modern Western 12 tone equal tempered scale. This would lead you to conclude that Balinese scales are wrong, in that some notes are a little too flat, and others a little too sharp. Similarly, you might conclude that Balinese rhythmic and musical structures are flawed and "primitive." But is this the right method for studying such music? In fact, Balinese musicians are highly accomplished, and have their own methods for teaching their music. They also have their own musical theory, according to which their scales, rhythms, and structures are correct; they do not orient to the twelfth root of two. (See [2] for a discussion of Balinese musical practice.) But in the nineteenth century and before, ethnocentric approaches were the norm, and non-Western culture was systematically devalued by such analyses. This paper suggests that similar things may be going on in much of today's Requirements Engineering.

Much traditional social science is based on a social scientist who stands outside the situation, using methods different from those used by the members of the culture to make sense of their world. To a great extent, this comes from the desire to be as "scientific" as the hard sciences, which are taken as prototypical of how to do science, combined with a fundamentally flawed understanding of how research is conducted in the hard sciences. The naive view of the hard sciences is that they achieve objectivity by banishing the experimenter from the experiment. But it is well known in quantum mechanics that measurements necessarily disturb systems, and it is also widely recognised in the philosophy of science that all measurements are necessarily made in the context of some theory, held by some theorist [22]. Thus the "method" of science that is used by traditional social science as a model does not hold even in the sciences that are taken to be exemplary.

This model of objectivity has always been dubious in social sciences such as anthropology and sociology in which participant observation is a key method. *Participant observation* is a method in which the observer attempts to become part of the community of interest, by developing a legitimate rôle within that community. For example, researchers have apprenticed as a mid-wife, jazz musician, waitress, etc. Recently, the post-modern movement in ethnography has studied the process of becoming a member, and the assumptions that underlie the belief that the ethnographer has become a member (e.g., [19]).

The assumption that social science methods differ from those used by the people studied is challenged by ethnomethodology, which argues that social scientists employ the same kinds sense-making activities as members of the culture studied [13]. This argues against scientific objectivity, or at least, against the claim that analysts have a unique access to objectivity.

6 Discourse Analysis

Within linguistics, the phrase "discourse analysis" is used most broadly to describe the study of structures larger than the sentence. This section describes both interactional and linguistic approaches to such structures. The interactional approaches arise from ethnomethodology, and illustrate how social order is reproduced in the particular but very important domain of conversation. The linguistic approaches arise from sociolinguistics, and concern the internal structure of certain discourse forms.

6.1 Conversation Analysis

Conversation Analysis grew out of ethnomethodology (see Section 5). It attempts to describe the underlying social organisation that makes conversation orderly and intelligible. Conversation is one of the most prevalent yet invisible forms of social interaction, and may be considered typical of how people construct their world in an orderly way [16].

Conversation is a folk term for activities that members might describe as sitting around and chatting, just talking, socialising, etc. However, as a technical term in

Conversation Analysis, conversation is that interactional system in which turns are not preallocated, i.e., in which the order of interaction is negotiated in real time, as the conversation procedes. By contrast, in forms of interaction such as debates, rituals, and seminars, the order of events, speakers, etc. is prearranged. For example, the order and the orderliness of a church service is not produced by the participants in the course of enacting it; there is no on-the-spot negotiation of whether the sermon shall precede or follow the collection.

6.1.1 Turntaking

Within conversation, turntaking is the basic system for creating social order. The order that it creates is the normative form of conversation: there should be one speaker at a time, with no gaps or overlaps [38]. It is important to note that what counts as a gap or overlap is culturally determined [41]. For example, what sounds like a long pause for a New Yorker may be barely noticeable for a New Englander.

In brief, to achieve turntaking, the current speaker speaks until he/she comes to a possible turn-transition place, i.e., a point which is semantically and syntactically a possible end of sentence. Then he/she may select another speaker, either verbally, by gesture, or by eye gaze, or another participant may self-select as the next speaker, or there may be a gap, i.e., a silence long enough in the particular culture to be noticed as such. The speaker may then continue, so that the possible between-turn gap becomes a within-turn gap. When there are overlaps, i.e., when two speakers speak at once, one drops out.

The important point is that turntaking is achieved in the moment by moment interaction of the participants. It is not the case that there are certain rules that define the set of all possible conversations; rather, the application of rules to particular situations is a matter of on-going work by the participants, who may, for example, negotiate the status of a particular silence.

6.1.2 Adjacency Pairs

While turntaking is an important part of the syntax of conversational organisation, adjacency pairs are a partially syntactic, partially semantic organisational structure. An *adjacency pair* is a pair (or larger set) of utterances "whose central characteristic is the rule that a current action (a 'first pair part' such as a greeting or question) requires the production of a reciprocal action (or 'second pair part') at the first possible opportunity after the completion of the first" ([16], p. 287).

Examples are sequences like question-answer and greeting-greeting, where one speaker's production of a question or greeting projects another speaker's production of an answer or second greeting.

Once a speaker has produced the first part of an adjacency pair, the second pair part can be *noticeably absent*. It is important to distinguish between an absence and a noticeable absence. At any point in a conversation, the range of things that are *not* said is infinite; but because the first pair part *projects* (or sets up the expectation of) the production of the second pair part, we can notice its absence. In fact, we as analysts can see speakers orienting to such an absence; for example, someone might say, "Don't you say hello?" in response to the absence of a second greeting. Such a response from a speaker shows that adjacency pairs are not merely a construct of the analyst, but in fact are categories that speakers themselves use to organise their conversations.

6.2 Discourse Units

Another approach within linguistics that is relevant for requirements elicitation is the study of the *discourse unit*, the linguistic unit directly above the sentence. Some very common examples of the discourse unit that have been studied extensively are the oral narrative of personal experience [21, 24, 35], the joke [37], the explanation [15], the spatial description [23, 27], and the plan [26]. As a structural unit, the discourse unit has two criterial properties: it has defined boundaries, and a describable internal structure.

The property of *definable boundaries* means that the discourse unit is a *bounded unit*; for example, with some interesting exceptions, we know when a speaker is or is not engaged in telling a narrative. Of course, there may be boundary disputes, either at the beginning, during which a speaker negotiates with hearers whether the narrative will be told, or at the end, where the speaker may negotiate the proper response to the unit with hearers [37, 34, 35]. However, such negotiations do not mean that the unit is not structurally bounded. Rather, they imply that the establishment of boundaries is a social construction, with serious social consequences for how the interaction can proceed.

One important effect of establishing of the boundaries of a discourse unit concerns turntaking. As we have seen, other things being equal, the sentence is the potential unit of turn exchange; i.e., a second speaker may begin to speak when the first speaker has reached a permissible end for his sentence. However, if the first speaker has negotiated permission to produce a recognised discourse unit, such as a joke or a story, then that speaker has the floor until the unit is completed. A second speaker may contribute questions, appreciations, side sequences, etc., but the discourse unit and topic in progress will not be changed until the unit is recognised as completed.

The second important property of the discourse unit is that it has a *precise internal structure* that is just as describable as sentential syntax. The description of this internal structure is necessary for understanding the interactional process of discourse construction, because the task of hearers is quite different, for example, in different sections of a narrative. Moreover, discourse structure can be described with just as much mathematical precision as

sentential syntax (see [26, 15] for some appropriate mathematical apparatus).

We expect that narratives will be particularly important for understanding the requirements process, because much of what is communicated between the parties will be framed as stories, e.g., about what our group does, what we hope to accomplish with the new system, what our problems are, etc. For example, a study of experienced photocopy repair personnel [33] shows that they often use narratives for informal training of novices in problems that are not covered in official manuals and training courses. These "war stories" are an important part of the work life of photocopy repair mechanics, although management may see this activity as 'goofing off' rather than as a legitimate part of the job. Also, [14] mentions a case study by the authors of this paper, in which evaluations extracted from jokes and stories were used to reconstruct a value system for an organisation, and where task oriented discourse was used to determine work structure.

6.3 Reproducing Social Order

We have discussed the orderly nature of social interaction, and indicated that this social order is produced by the participants in their moment to moment interactions. We have not yet considered how familiar social orders are reproduced: although participants are continuously producing social order, it always seems to be substantially the *same* order that is reproduced — the relations of class, gender, age, power etc., do not suddenly disappear, and are not suddenly produced in unusual or surprising ways. This observation is a necessary correction to a possible view of members' categories and members' methods which says that (for example) the structures of gender privilege, or of a ten ton truck bearing down upon you, are just your construction of the world, and if you don't like them, then you can just construct something else. Although few people will take such a naive constructivist attitude towards a truck, some do take it toward social structures, and thus the question must be explored.

There are material artifacts, histories of behaviour, interpretations of behaviour, social expectations of consequences, individual tastes and preferences, etc. that lead participants towards reproducing the same social order. For example, to illustrate the rôle of material artifacts in our example of the construction of a seminar, the social category of seminar is partly constructed by the turntaking behaviour of the participants. It is also constructed by the material artifacts and the ways in which people use them: the arrangement of a table in the room in a position that is understood to be the head, a board and writing materials that one participant uses and the others don't, perhaps a glass of water for one participant.

To illustrate the rôle of the interpretation of behaviour, we consider an example from turntaking, namely interruptions and overlaps, and their relation to gender rôles. A successful interruption is an example of a violation of a speaker's turn in which participant A begins to speak, participant B begins to speak while A is still speaking, and A then drops out. It has been found in U.S. data [52] that interruptions are very rare in same-sex conversations. In cross-sex conversations, from 75% to 90% of successful interruptions involve men interrupting women. (The percentages differ slightly, depending on the situation, and the degree of acquaintance of the conversational partners.)

Why is this? West and Zimmerman [48, 49] suggest that interruption by one's conversation partner is not only a consequence of lesser status, but is also a way of establishing and maintaining a status differential. For this formulation to make sense, it is necessary to understand in detail how participants in an interruption negotiate who is to drop out. When two participants start talking at once, or when one participant begins to speak while another is still speaking, one or both speakers may become louder, and continue to increase volume until the participant who is speaking more softly drops out. This appears to be a gender neutral description of the mechanism. However the social meaning of increasing volume is different for men and women. It is an indication of what kind of a person one is: in the case of men, a person who stands up for his rights, in the case of women, a strident and aggressive person. These different social meanings for the same behaviour ensure that it is almost always the woman who drops out of an overlap.

Some larger scale studies of social reproduction have considered class distinctions among adolescent school children in Britain and the United States [50, 11]. In each case, working class students' attitudes towards the importance of friendship networks and school culture exactly reproduced the kinds of behaviours, attitudes, preferences and skills that led to their being tracked to skilled or unskilled labouring jobs, rather than to higher education or managerial and professional work.

The reproduction of social order is an important issue for Requirements Engineering, because it is necessary to consider the effect of a new system on social structures, as suggested by the following questions: Will the new system reproduce the existing social order? Or will the order be altered in significant ways? Do the existing social structures suggest requirements that would negate the improvements expected from the new system?

7 Discussion

Every method has some limitations. Questionnaire-based interviews are limited by their stimulus-response model of interaction, which assumes that a given question (as stimulus) always has the same meaning to subjects. Moreover, this method excludes the kinds of interaction that could be used to establish shared meaning between the subject and the interviewer. Open ended interviews allow less

constrained interaction between the interviewer and the interviewee, who is no longer considered the subject of an experiment. However, this method is still limited by the need for the participants to share basic concepts and methods, without which they will be unable to negotiate shared meanings for the questions asked. Open ended interviews are also more vulnerable to distortion by interviewer bias. These limitations also apply to focus groups, and to their cousins in Requirements Engineering, JAD (or RAD) groups. In addition, these methods are vulnerable to political manipulations by participants. Protocol analysis involves an artificial discourse form, and is based on an incorrect cognitivist model of human thought that entirely ignores social context. None of these methods can elicit tacit knowledge, and all are subject to the say-do problem.

The principles of ethnomethodology, such as members' concepts and members' methods, provide a powerful framework for a deeper consideration of these limitations, and suggest that traditional sociology and its methods are based on faulty assumptions about how social interaction is organised.

Conversation, discourse and interaction analyses are only applicable to situations where there is significant social interaction; conversation and discourse analyses are only applicable to verbal data. But the most important limitation of these methods is that they are very labour intensive. In particular, it can take a highly skilled person a very long time to produce a transcript from a videotape of live interaction. Another limitation is that these methods cannot be (directly) applied to the study of systems that have not yet been built. However, they can be used to obtain tacit knowledge, because they bypass the unreliable explanations of users, and instead examine what they actually do.

Despite their limitations, we do not wish to suggest that any of these methods cannot be useful in requirements elicitation (with the possible exception of protocol analysis). In fact, their strengths seem to some extent complementary, so that combinations of the various methods can be usefully applied to particular problems. In particular, we suggest it is often a good idea to start with an ethnographic study to uncover basic aspects of social order, such as the basic category systems used by members, the division into social groups, the goals of various social groups, typical patterns of work, how current technology is used, etc. (see [36] for a review of ethnography in relation to Requirements Engineering). After this, one might use questionnaires or interviews to explore what problems members see as most important, how members place themselves in various classification schemes, etc. Then one might apply conversation, discourse or interaction analysis to get a deeper understanding of selected problematic aspects.

Techniques from discourse analysis can be useful when verbal communication is important to the system being developed; conversation analysis can also help to uncover limitations of other techniques. Some previous joint work of the authors, briefly described in [14], shows how the discourse analysis of stories can be used to explore the value system of an organisation, and how the discourse analysis of explanations can be used for a kind of situated task analysis. Interaction analysis can be used to discover details of non-verbal interaction in real work environments [18]; but the effort required to produce video transcripts suggests that this method should be used very selectively. Ethnography should be used continually to provide context for results obtained by other methods.

To sum up, we recommend a "*zooming*" method of requirements elicitation, whereby the more expensive but detailed methods are only employed selectively for problems that have been determined by other techniques to be especially important. From this point of view, the various techniques based on ethnomethodology can be seen as analoguous to an electron microscope: they provide an instrument that is very accurate and powerful, but that is also expensive, and requires careful preparation to ensure that the right thing is examined.

It is interesting to notice that all of these methods, including zooming, can be used not only for requirements elicitation, but also for studying the system development process itself, including the Requirements Engineering process. In this way, we may hope to develop a scientific methodology for systems development; in fact, we have already tried to do this in a limited way in this paper, by using concepts from ethnomethodology to explore the limitations of more traditional methods.

We close this paper with some research tasks that seem to merit further investigation:

1. Do detailed empirical studies of the entire system lifecycle, including the rôle of planning, management and phases, using ideas of Suchman [43]; in particular, investigate the hypotheses that requirements activities are distributed throughout the lifecycle, and that plans serve at least as much to justify actions as they do to predict them.

2. Do case studies to determine the rôle of political considerations in Requirements Engineering, and how they affect the use of various commercial methods and tools.

3. Do case studies to determine the limitations and strengths of JAD groups, in relation to the entire system lifecycle.

4. Work out detailed guidelines for the zoom method described above, and try it in some case studies. In particular, work out the relationships between discourse, conversation and interaction structures, and when each should be applied.

5. Do detailed empirical studies of the comparative effectiveness of various commercial methods and tools for various purposes.

We believe that if research projects along these lines were completed, then Requirements Engineering would be much closer to having a sound scientific foundation.

References

[1] Dorine C. Andrews. JAD: A crucial demension for rapid applications development. *Journal of Systems Management*, pages 23–31, March 1991.

[2] Judith Becker and Alton Becker. A musical icon: Power and meaning in Javanese gamelan music. In Wendy Steiner, editor, *The Sign in Music and Literature*, pages 203–228. University of Texas, 1981.

[3] Carl Bereiter and Siegfried Engelmann. *Teaching Disadvantaged Children in the Pre-school*. Prentice-Hall, 1966.

[4] Woodrow Wilson Bledsoe and Donald Loveland (editors). *Automated Theorem Proving: After 25 Years*. American Mathematical Society, 1984. Volume 29 of Contemporary Mathematics Series.

[5] Barry Boehm. *Software Engineering Economics*. Prentice-Hall, 1981.

[6] Richard Boland. In search of management accounting: Explorations of self and organization. Technical report, Case Western University, 1991.

[7] Francoise Brun-Cottan. Talk in the workplace: Occupational relevance. *Research on Language and Social Interaction*, 24:277–295, 1990/1991.

[8] Graham Button and Wes Sharrock. Occasioned practises in the work of implementing development methodologies, 1992. Draft, Rank Xerox EuroPARC.

[9] John Canny. *The Complexity of Robot Motion Planning*. MIT, 1988.

[10] William F. Clocksin and Christopher Mellish. *Programming in Prolog*. Springer, 1981.

[11] Penelope Eckert. *Jocks and Burnouts*. Teachers College, 1989.

[12] K. Anders Ericsson and Herbert A. Simon. *Protocol Analysis: Verbal Reports as Data*. MIT, 1984.

[13] Harold Garfinkel. *Studies in Ethnomethodology*. Prentice-Hall, 1967.

[14] Joseph Goguen. The dry and the wet. In Eckhard Falkenberg, Colette Rolland, and El-Sayed Nasr-El-Dein El-Sayed, editors, *Information Systems Concepts*, pages 1–17. Elsevier North-Holland, 1992. Proceedings, IFIP Working Group 8.1 Conference (Alexandria, Egypt); also, Programming research Group, Technical Monograph PRG–100, March 1992, Oxford.

[15] Joseph Goguen, James Weiner, and Charlotte Linde. Reasoning and natural explanation. *International Journal of Man-Machine Studies*, 19:521–559, 1983.

[16] Charles Goodwin and John Heritage. Conversation analysis. *Annual Review of Anthropology*, 19:283–307, 1990.

[17] Marina Jirotka. Ethnomethodology and requirements engineering. Technical report, Centre for Requirements and Foundations, Oxford University Computing Lab, 1991.

[18] Adam Kendon. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge University, 1990. Studies in Interactional Sociolinguistics Number 7.

[19] D.K. Kondo. *Crafting Selves: Power, Gender and Discourses of Identity in a Japanese Workplace*. Chicago, 1990.

[20] Gunter Kress and Robert Hodges. *Language and Ideology*. Kegan Paul, 1979.

[21] William Labov. The transformation of experience in narrative syntax. In *Language in the Inner City*, pages 354–396. University of Pennsylvania, 1972.

[22] Imre Lakatos. *The Methodology of Scientific Research Programmes*. Cambridge, 1978.

[23] Charlotte Linde. *The Linguistic Encoding of Spatial Information*. PhD thesis, Columbia University, 1974. Department of Linguistics.

[24] Charlotte Linde. Private stories in public discourse. *Poetics*, 15:183–202, 1986.

[25] Charlotte Linde. Personal experience, 1992. Institute for Research on Learning, Palo Alto CA.

[26] Charlotte Linde and Joseph Goguen. Structure of planning discourse. *Journal of Social and Biological Structures*, 1:219–251, 1978.

[27] Charlotte Linde and William Labov. Spatial networks as a site for the study of language and thought. *Language*, 51(4):924–939, 1975.

[28] Paul Luff, David Frohlich, and Nigel Gilbert, editors. *Computers and Conversation*. Academic, 1990.

[29] David L. Morgan. *Focus Groups as Qualitative Research*. Sage, 1988.

[30] Claus Moser and Graham Kalton. *Survey Methods in Social Investigation*. Gower, 1971.

[31] Allen Newell. Physical symbol systems. *Cognitive Science*, 4:135–183, 1980.

[32] Allen Newell and Herbert Simon. GPS, a program that simulates human thought. In Edward Feigenbaum and Julian Feldman, editors, *Computers and Thought*, pages 279–293. McGraw-Hill, 1963.

[33] Julian Orr. Narratives at work: Story telling as cooperative diagnostic activity. In *Proceedings, Conference on Computer Supported Cooperative Work (SIGCHI)*. Association for Computing Machinery, 1986.

[34] Livia Polanyi. So what's the point? *Semiotica*, 25(3–4):208–224, 1978.

[35] Livia Polanyi. *Telling the American Story*. MIT, 1989.

[36] Dave Randall, John Hughes, and Dan Shapiro. Steps toward a partnership: Ethnography and system design, 1992. Draft, Lancaster University.

[37] Harvey Sacks. An analysis of the course of a joke's telling in conversation. In Richard Baumann and Joel Scherzer, editors, *Explorations in the Ethnography of Speaking*, pages 337–353. Cambridge, 1974.

[38] Harvey Sacks, Emanuel Schegloff, and Gail Jefferson. A simplest systematics of the organization of turn-taking in conversation. *Language*, 504:696–735, 1974.

[39] Emmanuel Schegloff. Body torque, 1990. Paper presented at 89th Annual Meeting of the American Anthropological Association, New Orleans LA, November 1990.

[40] Ron Scollon and Suzanne B.K. Scollon. *Narrative, Literacy and Face in Interethnic Communication*. Ablex, 1981. The Literate Two-year-old: The Fictionalization of Self.

[41] Ron Scollon and Suzanne B.K. Scollon. *Narrative, Literacy and Face in Interethnic Communication*. Ablex, 1981. Chapter 2: Athabaskan-English Interethnic Communication, pp.11–38.

[42] David Stewart and Prem Shamdasani. *Focus Groups: Theory and Practice*. Sage, 1990.

[43] Lucy Suchman. *Plans and Situated Actions: The Problem of Human-machine Communication*. Cambridge University, 1987.

[44] Lucy Suchman. Constituting shared workspaces. In Yrjo Engestrom and David Middleton, editors, *Cognition and Communication at Work*. Cambridge, to appear.

[45] Lucy Suchman. Technologies of accountability: Of lizards and airplanes. In Graham Button, editor, *Technology in Working Order: Studies of Work, Interactioin and Technology*. Routledge, to appear.

[46] Lucy Suchman and Brigitte Jordan. Interactional troubles in face-to-face survey interviews. *Journal of the American Statistical Association*, 85(409):232–241, 1990.

[47] Francisco Varela, Evan Thompson, and Eleanor Rosch. *The Embodied Mind*. MIT, 1991.

[48] Candace West and Don H. Zimmerman. Women's place in everyday talk: Reflections on parent-child interaction. *Social Problems*, 24:521–529, 1977.

[49] Candace West and Don H. Zimmerman. Small insults: A study of interruptions in cross-sex conversations between unacquainted persons. In Barrie Thorne, Cheris Kramarae, and Nancy Henley, editors, *Language, Gender and Society*, pages 102–117. Newbury House (Rowley MA), 1983.

[50] Paul Willis. *Learning to Labor: How Working Class Kids Get Working Class Jobs*. Columbia University, 1977.

[51] Nessa Wolfson. *The Conversational Historical Present in American English Narrative*. Foris, 1982.

[52] Don H. Zimmerman and Candace West. Sex roles, interruptions and silences in conversations. In Barrie Thorne and Nancy Henley, editors, *Langauge and Sex: Difference and Dominance*. Newbury House (Rowley MA), 1975.