# Experiences from Conducting Semi-Structured Interviews in Empirical Software Engineering Research

**Siw Elisabeth Hove**
Simula Research Laboratory
Lysaker, Norway
siweh@simula.no

**Bente Anda**
Simula Research Laboratory
Lysaker, Norway
bentea@simula.no

## Abstract

*Many phenomena related to software development are qualitative in nature. Relevant measures of such phenomena are often collected using semi-structured interviews. Such interviews involve high costs, and the quality of the collected data is related to how the interviews are conducted. Careful planning and conducting of the interviews are therefore necessary, and experiences from interview studies in software engineering should consequently be collected and analyzed to provide advice to other researchers. We have brought together experiences from 12 software engineering studies, in which a total of 280 interviews were conducted. Four areas were particularly challenging when planning and conducting these interviews; estimating the necessary effort, ensuring that the interviewer had the needed skills, ensuring good interaction between interviewer and interviewees, and using the appropriate tools and project artifacts. The paper gives advice on how to handle these areas and suggests what information about the interviews should be included when reporting studies where interviews have been used in data collection. By sharing experience, knowledge about the accomplishments of software engineering interviews is increased and hence, measures of high quality can be achieved.*

## 1 INTRODUCTION

Qualitative research methods originate from sociology and anthropology [7], and were designed mostly by educational researchers and social scientists to study human behaviour [27]. Interviews are a frequently used technique for data collection within qualitative research.

Interviews are often used in empirical software engineering research. For example, in a case study that investigated the trade-off between inspections and testing in a software project, data from interviews was triangulated with observations and quantitative measures regarding faults and resources required to clarify the data collection, and to obtain the employees' opinions about the activities preformed [5]. Another example is an industrial case study in which semi-structured interviews with 16 system developers and project managers contributed to identifying difficulties with, and benefits, of applying UML-based development in large projects [1]. Further, an in-depth survey used interviews to obtain an overview of the estimation methods that software companies apply to estimate their projects, and to investigate why those particular methods are chosen, and how accurate they are [18].

The purpose of using interviews in empirical studies is often to collect data about phenomena that cannot be obtained using quantitative measures. In studies where the research goals are of a qualitative nature, it is appropriate to rely on qualitative measures. Interviewing people provides insight into their world; their opinions, thoughts and feelings.

However, interviews are a resource-demanding data collection method; activities such as planning, conducting and analyzing are time-consuming by nature. In addition, interviewees have to spend time on a "non-productive" activity. Furthermore, the interviews should be carried out carefully, because the way in which the interview is conducted determines the quality of the data collected. It is also important to ensure that the interviewees feel comfortable, so they are willing to share their experiences with the interviewer.

By systematically collecting and discussing interviewing experiences within the field, better studies can be conducted and thereby, the quality of the collected measures can be improved. In this paper, we present experiences from 280 interviews conducted in 12 software engineering studies in our research group at Simula Research Laboratory. We have identified four areas that deserve special attention when planning and conducting interviews: (i) issues related to the effort necessary to accomplish the interviews, (ii) the skills of the interviewer, (iii) interaction between interviewer and interviewees, and (iv) tools and project artifacts. These areas are broken down into subareas and discussed in more detail by combining our own experiences with existing literature on research interviews. Existing knowledge on how to conduct research interviews is thus adapted to software engineering. Factors related to analysis are not considered in this paper.

In our opinion, researchers reporting studies in which interviews have been used in the collection of data often fail to describe how the interviews were conducted. Based on our experiences, we suggest what information

about interviews and interviewees should be included when reporting such studies.

The remainder of this paper is organized as follows. Section 2 provides a brief overview of different types of interview and of a selection of the relevant literature. Section 3 reports the studies from which our experiences are collected. These experiences are discussed in Section 4. Section 5 suggests the information that should be reported about interviews. Section 6 concludes.

## 2 BACKGROUND

There exists a large amount of literature on research interviews, in which a variety of topics is covered. For this paper we have selected literature that either (1) deals with interviews in software engineering research, or (2) addresses one or several of the topics that we have identified as important for planning and conducting research interviews.

It is common to differentiate between *individual interviews*, in which an interviewer and a respondent talk about a theme of common interest [15], and *group interviews*, also known as focus groups [20], in which several subjects discuss topics introduced by one or two interviewers.

Furthermore, it is common to distinguish between structured and unstructured interviews. In *structured interviews*, the interviewer has very specific objectives for the type of information sought for in the interview, so the questions can be very specific [26]. In highly structured interviews, all the answers can be quantified. In *unstructured interviews*, in contrast, the interviewer suggests the theme for the interview, but has few specific questions in mind [25]. Many studies employ a combination of these two types; *semi-structured interviews* [25], also called focused interviews [17]. Such interviews combine specific questions (to bring forth the foreseen information), and open-ended questions (to elicit unexpected types of information).

Despite the fact that interviews are used frequently in software engineering, few papers address factors related to the planning and conducting of the interviews. One exception is a paper by Seaman [26], which outlines different objectives for conducting interviews: collecting historical data, eliciting opinions and impressions, identifying terminology, and clarifying things that happened during the study in cases where interviews are used in combination with observations. Furthermore, it addresses how much information the interviewer should give the interviewees about the study. Techniques for handling subjects that either say the barest minimum or stray from the topic of the question are outlined. The contents and purposes of an interview guide are described and different ways to record the interviews, such as using audiotape or a scribe, are presented.

Several sources within qualitative research provide information about interviewing. Opinions on how much the interviewers need to know about the topics under study are provided in [15, 8]. In [15], the importance of the interviewer being an expert of the topic of the interview is outlined, while in [8] it is claimed that it is easier for researchers to study areas in which they have little or no experience.

Desired interviewer skills are outlined in [15, 27, 24]. Such skills involve being nonjudgmental and sensitive, letting people talk, and paying attention and express the questions clearly. Both [28, 26] address issues related to problem respondents. Different types of respondents, such as unresponsive respondents, are described in [28].

How to approach sensitive questions is outlined in [15, 8, 25], and ways to create an atmosphere of trust are suggested, e.g., ensuring confidentiality, phrasing questions in a non-threatening manner and not express dismay or openly disagree with what the respondent says.

Different types of interview questions are described in [15, 27, 21, 24], such as experience questions, opinion questions and feeling questions. It is recommended that interviewers avoid "why"-questions and questions to which the answer can only be "yes" or "no" [15, 21].

Tape recorder usage is frequently discussed. Many researchers use a tape recorder to avoid loss of information. The tape recorder also permits the interviewer to be more attentive to the subject, as outlined in [23, 28, 27, 21]. However, there are situations in which tape recording is inappropriate [25].

Different elicitation techniques for interviewing are described in [8]. Visual elicitation, where pictures and drawings etc. can be used to help interviewees to remember, is explained.

## 3 EXPERIENCE COLLECTION

Table 1 gives an overview of the studies and interviews upon which this paper is based. The authors of this paper were responsible for the interviews in four of the studies, including the DES study, which is the largest study with interviews conducted by our research group. We collected opinions from our colleagues regarding the other eight studies.

In what follows, examples from the DES case study will be used to illustrate several of the areas discussed. The experiences obtained from the other studies are referred to in more general terms. When speaking of "our experience", we mean the experiences of both the authors and our fellow researchers.

**Table 1. Overview of studies**

| Topic of study | Num interviews | Num subjects | Interview goal | Related papers |
|---|---|---|---|---|
| Effects of effort on front-end activities (DES) | 98 | 13 | Collect experiences of developers and managers through the planning and development phases. | In preparation |

| | | | | |
|---|---|---|---|---|
| Visualization in schema evolution | 9 | 1 | Study usability and functionality of a software tool | [13] |
| Schema evolution in industry | 3 | 3 | Locate problems in software engineering companies and analyze impact of change | [12] |
| Estimation survey | 70 | 70 | Study estimation practices in industry | [19] [20] |
| UML & Feedback Collection | 19 | 19 | Investigate the use and impact of UML documentation<br>Gather opinions using a feedback collection tool | [4] [14] |
| Estimation Accuracy | 6 | 6 | Study estimation accuracy in context of client-supplier relationships. | In preparation |
| Changeability in evolutionary development | 6 | 6 | Study end-users, documentation and technology risks, and how these factors influence changeability in evolutionary development | [3] [16] |
| Estimation experiment | 27 | 27 | Collect information about estimation practices and evaluate think-aloud in an experiment studying task level effort estimation | [6] |
| Use case based estimation | 11 | 11 | Evaluate and improve a method for use case based estimation in a company. | [2] |
| UML-based development | 16 | 16 | Evaluate and improve UML-based development in a company. | [1] |
| Estimation Impact | 7 | 7 | Understand how effort estimates affect project work | [10] |
| Estimation check list | 8 | 8 | Create a check-list for estimation | [11] |
| **Total** | **280** | **187** | | |

## 4 INTERVIEWING EXPERIENCES

This section reports our experiences with planning and conducting interviews regarding effort, skills of the interviewer, interaction issues, and appropriate tools and project artifacts.

### 4.1 Necessary Effort

Understanding how much effort an interview study requires is important in the planning phase for allocating resources and scheduling the study. In what follows, the different activities related to research interviews are described. Furthermore, whether it is better to have one or two interviewers is also discussed.

#### *Activities*

Interview studies include several activities that require effort, in addition to the actual conduct of the interviews. Time estimates are given on two of the activities: summary writing and transcriptions. It is impossible to give time estimates for all the activities because of differences between the studies. Nevertheless, by being aware of these activities, it is more likely that realistic plans and estimates can be made. In our experience, the activities that are required in an interview study, in addition to the actual interview, are as follows:

- **Scheduling** It is necessary to make appointments with interviewees. If it is necessary to recruit or select subjects carefully, this activity will probably require a lot of effort.
- **Collecting of background information** In some of the studies that we conducted, it was necessary for the researchers to read project documentation, to collect information from personal CVs, etc.
- **Preparing interview guides** The time required for preparing interview guides varied substantially from study to study. Some factors can increase the time and effort required, for example, several researchers making the guide together or needing to adapt the guides to each subject.
- **Discussions/meetings** If several researchers are involved in the study, it might be necessary to spend time on meetings and discussions before, after or between the interviews.
- **Summary writing** In our experience, it takes approximately three to four hours to summarize one hour of audio taped interview.
- **Transcribing** It normally takes about eight hours to transcribe one hour of audio taped material.

Effort in relation to analysis must also be taken in to account. Analysis is time consuming in nature and will probably exceed the time spent on the activities listed above by several times.

#### *One or two interviewers*

A research interview is usually conducted by one interviewer. However, in some situations, it can be beneficial for two interviewers to conduct the interview together.

In the DES case study, two interviewers had different roles; one led the interview, while the other took notes and asked additional questions when it was appropriate. The general impression was that (i) the interviewees talked more with two interviewers than with one and that (ii) more follow-up questions were asked in the two interviewer case, because two interviewers often had more input

and ideas than one interviewer. To explore this observation further, we compared how long the interviewees talked when one interviewer conducted the interviews with how long they talked with two interviewers (Table 2). In a qualitative study, all data is potentially useful, so it is assumed that everything said in the interviews was valuable and relevant.

**Table 2. Interview length with one and two interviewers**

| Subject No | One interviewer | Two interviewers | Increase (percent) |
|---|---|---|---|
| **1** | 19 | 35 | 84,94 % |
| **2** | 17 | 38 | 122,29 % |
| **3** | 29 | 27 | -5,57 % |
| **4** | 29 | 32 | 11,58 % |
| **5** | 32 | 33 | 3,41 % |
| **6** | 24 | 44 | 80,46 % |
| **7** | 12 | 43 | 239,54 % |
| **8** | 41 | 38 | -7,73 % |
| **9** | 27 | 46 | 71,24 % |
| **10**[1] | 26 | 25 | -3,02 % |
| Mean | 26 | 36 | 59,71 % |

The numbers appearing in the columns named "One interviewer" and "Two interviewers" show the average length of interviews (in minutes) conducted, respectively, by one interviewer or two. The data is based on an equal numbers of interviews for each subject, i.e., the data for subject 1 is based on three interviews of each type. The numbers appearing in the column named "Increase (percent)" show in percent how much longer the interviews lasted with two interviewers.

The "increase" column shows that half of the subjects talked much more when the interviews were conducted by two interviewers. The other half had minor differences in interview length between the two interview types. This indicates that it might be preferable to have two interviewers, with respect to the amount the interviewees talk. It is assumed that the more the interviewees speak, the more information they provide, but the results must be interpreted with care, due to the small number of data points.

Based on these results from DES and experiences from other studies, we have identified the advantages and disadvantages of having two interviewers rather than one. We claim that the advantages of having two interviewers are as follows:

- In most situations, two interviewers will ask more questions than one interviewer. This will lead to the subjects talking more and thus, more information will be collected.
- It is often easier to use two interviewers than one because of the possibility of dividing the responsibilities. In semi-structured interviews, the interviewer must sometimes improvise, and it is challenging to listen closely to what is said and at the same time plan the next question. Thus, is beneficial to have a second interviewer who can focus on what is said, ask follow-up questions and aid the primary interviewer when necessary.
- Two interviewers have the opportunity to discuss and verify their interpretation of the interview. This increases the probability of understanding the subject correctly.

However, there are also disadvantages to having two interviews. The workload associated with some of the interview activities will be doubled, and it requires more planning to conduct the interview with two interviewers than with one. The researchers must define their roles and agree upon the structure and content of the interview so they do not pull in different directions.

### 4.2 Qualifications

It is obvious that the qualifications of the interviewer will have a great effect on the quality of the interview and thereby, also the quality of the collected data. In software engineering interviews, it is important that the interviewer has comprehensive knowledge of the software engineering area under investigation, in addition to good interviewing skills.

#### *Interviewing skills*

Good interviewing skills include the following:

- encouraging the interviewees to talk freely,
- asking relevant and insightful questions, and
- following up and exploring interesting topics.

The desired skills of the interviewer are explored and discussed thoroughly in the qualitative interview literature. According to [15], an interviewer must master human interaction and have excellent conversational skills. The interviewer should open the interview by re-explaining confidentiality and anonymity [23], and explain the purpose of the interview and his or her intentions [23, 15]. It is also important to ask for permission in case the interview is being audio recorded [28, 23]. In order to create a comfortable atmosphere in which useful information can be elicited, the interviewer should be nonjudgmental and sensitive, let people talk, and pay attention, as outlined in [27]. These desired qualities are also found in [15], in which it is also emphasized that the interviewer must express him- or herself clearly, be gentle, open and sensitive towards the interviewees, but at the same time not take everything said at face value and bear in mind what a subject has said ear-

[1] Three of the subjects in the study were only interviewed with one interviewer and are not included in this table.

lier in the interview. It is also important to clarify issues during the interview [15]. Further, the interviewer should enjoy the interview, or at least give the appearance of doing so [24]. These skills are developed mainly through practice. Pilot interviews and role plays should be conducted to train the interviewer [15].

The advice found in the literature is useful and in accordance with our experience. In the following, we provide some practical advice and a number of concrete examples that complement and re-emphasize the recommendations found in literature:

- Use a tidy and well-organised process: schedule the appointments as early as possible and be flexible to changes.
- Gain the trust of the subject: ensure confidentiality, explain your research motives, describe how the data will be applied and in which papers they are going to be used.
- Be courteous at all times. Remember to thank the subjects for their contribution.
- Be well prepared. Bring slides to the interview with the interviewee's name on, your research goals, a short presentation of the research institution you represent etc.
- Take care with your appearance and make a good first impression. We have experience with two strategies; 1) dressing up to make a formal and serious impression, 2) dressing down to appear less threatening. A good rule of thumb is to dress at approximately the same level of smartness and formality as the interviewees. If the interview is conducted in the subjects' work place, it is beneficial to have an idea of the dress code in the company and adapt to it.
- Allow the interviewees to view the questions in advance, so they can prepare for the interview.
- Talk informally with the interviewees before or after the interviews to facilitate a friendly and relaxed atmosphere.
- Use humour. This can contribute positively to the interview. Laughing together can create a more relaxed and open climate. However, humour and bonhomie must not be used excessively or in inappropriate situations, because it is important to project an image of gravitas and seriousness.
- Be active and show interest by nodding, paying attention and asking follow-up questions.
- Be careful not to argue or question the answers you get. The interviewee may become defensive and lose respect for you.

### *Qualifications in software engineering*

In [15] it is claimed that a good interviewer is an expert on the topic of the interview. The interviewer should have extensive knowledge of the theme of the interview so he or she can conduct an informed conversation and know which topics it is important to follow up. The interviewer will then be able to steer the interview so that useful information is obtained.

A different view regarding the knowledge of the interviewer is presented in [8]. There, it is claimed that it is easier for researchers to study areas in which they have little or no experience, because familiarity with a topic might cause them to overlook details about which they assume they are already informed. Further, researchers familiar with the topics of study may give the informants the impression that they already know the answers and consequently, the informants may feel that the researchers are testing them.

Most of the interviews shown in Table 1 were conducted by researchers with substantial knowledge and experience of the topic of the interviews. However, some were conducted by interviewers with less knowledge of the studied domain. Our experience is in accordance with [15]. In the field of software engineering it is very important for the interviewer to have extensive knowledge of the theme of the interview. Armed with such knowledge, the interviewer will be able to understand the information that the interviewees give, ask the relevant follow up questions, and be able to clarify ambiguities and to control the interview. It can be difficult for an interviewer with limited knowledge to understand what is important and to follow up interesting and relevant topics further. If the interviewer does not have appropriate knowledge, the quality of the interview depends to an overwhelming extent on the interviewee.

## 4.3 Interaction issues

This section discusses experiences related to interaction between the interviewer and the interviewees. Issues such as how to handle a subject that either says the barest minimum or strays from the question are outlined. Our experiences with interview questions and in particular sensitive questions are covered.

### *Silent interviewees*

Sometimes the interviewer faces subjects that barely talk. Their answers are short and they are unwilling to elaborate. For interviewers, such subjects can be challenging.

Unresponsive subjects are described in [28]. Subjects may be unresponsive because they fear the consequences of talking too much, or see no potential benefit in participating in the interview and therefore no point in cooperation. According to [28], it can be very difficult to obtain information from them.

Strategies to cope with subjects that say the barest minimum are suggested in [26], such as asking questions

that cannot be answered only with "yes" or "no". Furthermore, the interviewer can feign ignorance and ask for details that are already well known to the interviewer. It is important not to be perceived as being too much of an expert. In addition, it is important to ensure that the interviewees understand that there are no "right" answers, because software developers tend to believe that anyone coming to interview them is really there to evaluate them.

In the interviews shown in Table 1, we have met two different types of silent interviewee: the uninterested ones, as described in [28] and the shy ones that are often very technically orientated.

Regarding *unresponsive interviewees*, the reason for their unresponsiveness can vary. Sometimes, their managers have told them to participate and consequently, they are not motivated to proffer information in the interview. In other cases, the interviewees do not believe that the interview is relevant to what they are doing.

In general, our experience is that it can be difficult to make these subjects more responsive. In cases where the interviewees are told to participate, one possible approach is to motivate them in advance; to contact them directly and tell them about your goals and intentions. Another possible strategy in such cases is to have two interviewers. For example, in the DES case study, the subjects that could be classified as unmotivated talked much more when two interviewers were present (e.g., subject 7).

In cases where the subjects do not find the interview relevant, one strategy for dealing with the problem is to understand why the interviewees feel this way and try to adapt the questions so that they become more relevant to them.

With *shy interviewees*, the approach should be different. The interviewer must make the interviewees feel comfortable. In the DES study, the shy respondents were also very technically orientated. An approach that worked well was to warm them up by encouraging them to talk about issues that they were really interested in, such as technical issues, although this was not very relevant for the research questions. Consequently, it was easier to make them talk about relevant topics later in the interview. Similarly, it is important to consciously use general interviewing skills, as mentioned in section 4.2, such as talking informally and not appearing threatening, to create a comfortable climate for the shy interviewees.

### *Interviewees that talk too much*

It is often easy to interview subjects that are very verbal and provide much information. Yet, there are subjects that talk too much about irrelevant topics.

This issue is discussed in [26]. There, it is claimed that on the one hand, interview time is valuable and should not be wasted, while on the other hand, in qualitative studies all the data is potentially useful and which data is actually useful can only be understood after it has been collected. It may therefore be better to let the interviewee ramble, since the rambling might make sense in hindsight. Steering a talkative subject back to the topic must be done gently.

The necessity of stopping a highly verbal respondent who goes off track is emphasized in [21]. The first step in stopping these respondents can be to give cues that discourage talking, such as stopping nodding the head, interjecting a new question as soon as the respondent pauses for breath and stopping taking notes. If these tactics are not working, it becomes necessary to interrupt. Although this might feel awkward and impolite, it is argued in [21] that it is both patronizing and disrespectful to let the respondent run on when no attention is being paid to what is said.

When very talkative subjects go off track, we have positive experiences with letting them finish and then changing the topic. We agree with [26] that it might be better to err on the side of letting them talk too much. During an interview, it can be difficult to know whether or not the information will be found to be useful during the analysis. Cutting the subjects off too soon might result in the loss of relevant information. However, if it is evident that the information is not relevant, or that there are time constraints, it will be necessary to interrupt them tactfully.

### *Interview questions*

This section addresses interview questions. Firstly, an overview of types of question described in the literature is provided. Secondly, our experience with different types of question in terms of how good we perceived them to be is detailed. The same is done with question techniques.

#### *Types of questions*

In [21] six different question types are explained; 1) Behaviour/experience questions, which elicit descriptions of experiences, behaviour and actions, 2) Opinion/value questions, which investigate what people think about certain issues, 3) Feeling questions, which aim at understanding the emotional responses of people to their experiences and thoughts, 4) Knowledge questions, which identify what factual information the respondent has, 5) Sensory questions, which capture the experiences of the senses, and 6) Background/demographic questions, which identify the characteristics of the person being interviewed. All of these questions can be asked in the present, past or future tense. Similar types of questions are outlined in [15, 27]. Both [15] and [21] state that "what" and "how"-questions should be asked, but "why"-questions and questions to which only "yes" and "no" answers are possible should be avoided. The use of leading questions is discussed in [15] and it is argued that they might enhance the reliability of the interviews.

Our experiences with different types of question are as follows:

- Questions where the interviewees must describe how they work often provide rich information, i.e., "Can you describe what you have been working with?" or "Can you explain how you tested the solution?" This corresponds to the behaviour/experience questions described in [21].
- Reflexive questions also provide useful information, such as "What could have been done differently in the project?" or "What was most challenging for you as a developer/manager in this project?" This corresponds to the opinions/value questions described in [21].
- Very detailed questions, such as "How many lines of code have you written in your career?" can be difficult to answer.
- Series of questions that presuppose that some activity has been completed successfully can be risky. If the activity has not been completed successfully the questions are without value. To illustrate, if it is presupposed that participants in the experiment used UML documentation and the interviewer prepares many questions in relation to this, those questions will be useless if the participants did not use such documentation at all.

### *Question techniques*

Question techniques such as probes and prompts are described in [24]. A probe is a device to get the interviewee to expand on a response. An obvious strategy is asking "Anything more?", but one can also, for example, use a period of silence or an enquiring glance to encourage more information. Prompts suggest to the interviewer the set of possible answers that the interviewer expects. Probes and follow up questions are also described in [15, 27, 22].

Laddered question techniques [22] enable the interviewer to consider questions in terms of how intrusive they are, e.g., action questions are less intrusive than feeling questions. By studying the subject closely and adjusting the level of intrusion by choosing the right types of question, quality interviews can be conducted.

In addition to use probes and asking follow up question, we have also applied the following techniques:

- It is valuable to ask informal questions at the beginning, in order to loosen up and create a relaxed atmosphere.
- The interviewees' ability to answer very general questions satisfactorily varies to a large extent. To approach this, it is possible to either 1) ask very specific questions at the beginning, such as "Can you describe how you worked with this task?", and then ask more general follow-up questions, such as "Is this how you normally work?" or 2) start with general questions and follow up with more specific questions to ensure that the important issues are covered.

### *Sensitive questions*

Questions that touch upon sensitive issues require special attention. Individuals might be uncomfortable with certain questions and hence might be unwilling to discuss certain topics, might hold back information or might not be completely honest. It is important to avoid situations in which subjects feel that the questions are so intrusive that they are uncomfortable in the interview situation. It can be devastating for a study if some of the important subjects refuse to participate in the interview because they feel offended or that their privacy has been invaded. It is therefore important to handle such issues with care.

Dealing with sensitive question is a topic that has received significant attention in literature. When approaching sensitive topics, it is important to create an atmosphere of trust and ensure confidentiality [15]. Furthermore, sensitive questions should be asked late in the interview, after the researcher has indicated that he or she understands the subject and is sympathetic to the interviewee [25]. Phrasing of the question is also considered important: words must be chosen carefully, and questions should be phrased in a general manner to avoid the feeling of personal exposure. Using words such as "we", "us", "they", and "them" can ensure that the researcher is on the same side as the respondent. Furthermore, a common technique involves "normalizing perceived deviance" [9], which entails that the interviewer should not express dismay or openly disagree with what the respondent says. Hence, the interviewees do not feel judged and an open atmosphere is created.

Our studies have led us to realise that questions related to the following areas are potentially sensitive in software engineering studies:

- issues related to the economy
- opinions about colleagues and customers,
- explanations of why things went wrong
- questions related to the interviewee's own competence and mistakes.

When approaching potentially sensitive questions, we have been careful to follow many of the guidelines mentioned in the existing literature, such as ensuring confidentiality, asking the sensitive questions late in the interview, phrasing the questions in a polite and respectful manner, avoiding the most sensitive areas and refraining from expressing dismay at what the respondent says.

We have noted that when these precautions are taken, the interviewees are surprisingly open and informative on sensitive issues. It seems that people involved in software development like to talk to an external party about such topics, provided that they believe that they will not be judged.

Questions in software engineering are more impersonal and less emotional than in other disciplines, such as psychology or sociology. In addition to the precautions taken, this can explain why the interviewees do not seem to

mind answering sensitive questions. However, it might be that the ease with which the respondents answer potentially sensitive questions is culturally dependent. Our experiences are mainly from the Norwegian IT industry and it is possible that this observation cannot be generalized to other countries.

### 4.4 Tools and artifacts

It is important to ensure that the appropriate tools and artifacts are used in interviews. This section discusses audio taping and illustrates how project artifacts can help in eliciting useful information.

#### *Tape recorder*

It is useful to keep a record of the interviews so the analysis can be based on accurate renditions of what was said. Video taping, audio taping and note taking during the interview are commonly used techniques. Which one to choose depends on the level of detail needed, the degree to which recording equipment disturbs the interviewee, and the interviewers' preferences. The most common method is to use a tape recorder. The advantages and disadvantages of tape recorder usage are discussed in [25]. On the one hand, audio recording the interviews helps to get the material written down in an accurate and retrievable form and additionally, the interviewer can concentrate on what is being said instead of also focusing on taking notes. On the other hand, the interviewees can feel uncomfortable with it. Furthermore, the recording equipment requires the interviewers' attention to ensure that it works properly. Finally, a lot of work with transcriptions of the tapes must be taken into account.

Contrary to [25], we have never met interviewees who have opposed the use of a tape recorder or have become uneasy in its presence. The reason for this can be related to the section discussing sensitive questions in interviews; it is possible that interviewees are comfortable talking about topics related to software development and therefore do not mind that the conversation is audio taped. However, as for sensitive questions, this can be culturally dependent and may not apply to all software engineers.

Regarding taking notes instead of audio recording the interviews, our experiences vary to some degree. In the DES case study, we both took notes and audio recorded all the interviews. In this case it was evident that the notes were not sufficient to capture all the details of the interview. Furthermore, when listening through the tapes, we realized that it is easy to remember things slightly differently from what was actually said. From other studies, we have similar experiences with audio recorder usage increasing the detail richness and have come to realise that much information is lost otherwise. However, in some of the studies presented in Table 1, the interviewer was very skilled in typing and was therefore able to take notes on the computer, have eye contact with the interviewees and be focused on the conversation. In such cases, the need for a tape recorder is smaller. However, a number of interviewers felt that if the interviewer is less skilled in typing and prefers to focus at one thing at a time, taking notes can be distracting and leads to an impaired ability to focus on the conversation.

The level of detail necessary for the study determines how important it is to record the interview. However, since none of us have been in situations where the tape recorder has been perceived as a disadvantage, we claim that it is advantageous to use it, at least as a backup. Although some are capable of taking good notes, it is better to have too much information than too little.

To round off this section; we strongly recommend the use of recording equipments that smooth the process of transcribing or making summaries. For example, a digital tape recorder with transcription software and a foot pedal to control the sound is ideal for processing the recordings easily. With unsuitable equipment, the time required for transcription and summary writing can increase substantially.

#### *Visual artifacts*

Sometimes it can be difficult to elicit information about software development just by asking verbal questions. The informants' knowledge may be tacit and difficult to explain [8]. During interviews, it is common to use different elicitation techniques to encourage the informants to reveal what they know, feel, think or believe. Visual techniques can be more appropriate than verbal elicitation in certain situations. Photographs, drawings, artifacts or items can be used. Such visual stimuli are referred to as projective aids or devices [8]. The interviewer asks the informants to describe what they see, explain the visual item, compare items or express their feelings about the item.

Visual elicitation was used in some of the interviews in the DES case study. Since we had access to all project documents, we brought artifacts such as UML diagrams, screenshots and other project documents to the interviews. The interviewees were asked to explain how they used the artifacts in their work and how they planned to use the items further along in the development process. We observed that using visual elicitation techniques had the following positive effects:

- It became easier for the interviewees to remember what they had been working with.
- It became easier for the interviewers to ask good follow-up questions related to the artifacts.
- The interviewees talked a great deal and provided rich and informative information.

In a few of the other studies reported in Table 1, project items, such as UML diagrams and code excerpts, were also applied with the same positive experiences as in the DES case study. Although visual elicitation is perceived as a helpful tool in some studies, this technique is not suitable for, or possible in, all studies.

## 5 REPORTING FROM INTERVIEWS

The quality of the data obtained from semi-structured interviews depends on the planning and conducting of the interviews. In order to assess results obtained through interviews, it is therefore important that sufficient information about the interviews is reported. In our opinion, the following information is the minimum that should be described in studies reporting research interviews:

- The interviewees should be described in terms of number of interviewees, how they were selected (e.g. their roles in the project or company under study) and how they were recruited.
- The interviews should be described in terms of number of interviews (possibly number of interviews with each interviewee) as well as duration and location of the interviews.
- The number of interviewers and their roles if several.
- The interview guide(s) that were used during the interviews, as well as other tools or project artifacts.

## 6 CONCLUSIONS

In this paper, we have shown that semi-structured interviews are frequently used as a data collection technique within the field of software engineering. Semi-structured interviews involve high costs, and the quality of the collected data is related to how the interviews are conducted. In addition, it may be challenging to ensure that the interviewees experience the interview in a positive way.

Consequently, we believe that advice on planning and conducting interviews may be useful for many researchers planning to undertake research involving interviews. It is further important to collect, systematize and share experiences with interviews within the field of software engineering in order to increase the probability of collecting measures of high quality. Based on experiences from 280 interviews, we have identified four main areas that are central when planning and conducting interviews: (i) estimating the necessary effort, (ii) ensuring that the interviewer has the required skills, (iii) ensuring good interaction between interviewer and interviewees, and (iv) using the appropriate tools and project artifacts.

Activities that require a great deal of time in addition to the interview are identified, e.g., developing interview guides, scheduling of interviewees, and transcribing interviews. Furthermore, we have reported the advantages and disadvantages of having one or two interviewers. Advice on how to improve interview skills is provided, and it is argued that it is very important that the interviewer is knowledgeable in the field. Experiences of dealing with difficult interviewees are also provided. Our experiences with different types of questions are described. Particular focus was put on questions that may be sensitive for the interviewees. We argue that it is important to be cautious when asking such questions, but also that people involved in software development generally are open about such issues. We further claim that it is preferable to audiotape interviews to avoid loss of information. Project artifacts, such as UML diagrams, code and other visual items can be employed successfully, to make remembering and talking easier.


## ACKNOWLEDGEMENT

We wish to thank Erik Arisholm, Stein Grimstad, Magne Jørgensen, Vigdis By Kampenes, Amela Karahasanovic and Kjetil Moløkken-Østvold for contributing to this paper by sharing their interviewing experiences. We also acknowledge Dag Sjøberg and Chris Wright for valuable comments on this paper.


## REFERENCES


[1] Anda, B., Hansen, K., Gullesen, I., and Thorsen, H. *Experiences from using a UML-based development method in a large organization.* Accepted for publication in Empirical Software Engineering, 2005.

[2] Anda, B., Angelvik, E., and Ribu, K. *Improving Estimation Practices by Applying Use Case Models. 4th International Conference on Product Focused Software Process Improvement (PROFES 2002), Rovaniemi, Finland*, 2002. LNCS 2559, Springer-Verlag, pp. 383-397.

[3] Arisholm, E. *Empirical Assessment of Changeability in Object-Oriented Software.* 2001. PhD thesis, Department of Informatics, University of Oslo.

[4] Arisholm, E., Ali, S., and Hove, S. An Initial Controlled Experiment to Evaluate the Effect of UML Design Documentation on the Maintainability of Object-Oriented Software in a Realistic Programming Environment. 2003. Simula Research Laboratory Technical Report 2003-04.

[5] Berling, T., and Thelin, T. *An Industrial Case Study of the Verification and Validation Activities. Proceedings of the Ninth International Software Metrics Symposium (METRICS 2003).* 2003. Sydney, Australia. pp. 226-238.

[6] Bratthall, L., Arisholm, E., and Jørgensen, M. *Program Understanding Behavior during Estimation of Enhancement on Small Java Programs. 3rd International Conference on Product Focused Software Process Improvement (PROFES 2001).* 2001. Kaiserslautern, Germany.

[7] Denzin, N., and Lincoln, Y. *Handbook of qualitative research.* 2000. Sage Publications: Thousand Oaks, CA.

[8] Johnson, J., and Weller, S. Elicitation Techniques for interviewing, In: *Handbook of interview research: context & method,* 2002. Gubrium, J, and Holstein, J. (eds), Sage Publication: Thousand Oaks, CA.

[9] Johnson, J. In-depth Interviewing, In: *Handbook of interview research: context & method*, 2002. Gubrium, J, and Holstein, J. (eds), Sage Publication: Thousand Oaks, CA.

[10] Jørgensen, M., and Sjøberg, D. *Impact of Effort Estimates on Software Project Work*, 2001. Information and Software Technology, 43, pp 939-948.

[11] Jørgensen, M., and Moløkken-Østvold, K. *A Preliminary Checklist for Software Cost Management. IEEE International Conference on Quality Software*, 2003. Dallas, USA, pp. 134-140.

[12] Karahasanovic, A. *Supporting Application Consistency in Evolving Object-Oriented Systems by Impact Analysis and Visualization*. PhD thesis, University of Oslo, 2002.

[13] Karahasanovic, A., and Sjøberg, D. *Visualizing Impacts of Change in Evolving Object-Oriented Systems: An Explorative Study. Proceedings of the International Workshop on Graph-Based Tools.* 2002. Barcelona, Spain. pp. 22-31.

[14] Karahasanovic, A., Anda, B., Arisholm, E., Hove, S., Jørgensen, M., Sjøberg, D., and Welland, R. *Collecting Feedback during Software Engineering Experiments*. Accepted for publication in Empirical Software Engineering, 2005.

[15] Kvale, S. *InterViews: an introduction to qualitative research interviewing*, 1996. Sage Publications: Thousand Oaks, CA.

[16] Lien, A., and Arisholm, E. *Evolutionary Development of Web-applications - Lessons learned, European Software Process Improvement Conference (EuroSPI 2001)* 2001. Limerick, Ireland.

[17] Merton, R., Fiske, M., and Kendal, P. *The focused interview: A manual of Problems and Procedures.* 1990. (2nd ed.) Free Press, NY.

[18] Moløkken-Østvold, K., Jørgensen, M., Tanilkan, S., Gallis, H., Lien, A., and Hove, S. *A survey on Software Estimation in Norwegian Industry. In 10th International Symposium on Software Metrics (METRICS 2004)*, 2004. Chicago, Illinois, USA: IEEE Computer Society, pp. 208-219.

[19] Moløkken-Østvold, K., Lien, A., Jørgensen, M., Tanilkan, S., Gallis, H., and Hove, S. *Does Use of Development Model Affect Estimation Accuracy and Bias? Product Focused Software Process Improvement: 5th International Conference (PROFES 2004)*, 2004. Kansai Science City, Japan, pp. 17-29.

[20] Morgan, D. *Focus groups as qualitative research*, 1997. Sage Publications: Thousand Oaks, CA.

[21] Patton, M. *Qualitative evaluation and research methods*, 1990. Sage Publications: Newbury Park, CA.

[22] Price, B. *Laddered questions and qualitative data research Interviews*, Journal of Advanced Nursing, 2002.

[23] Rapley, T. Interviews, In: Seale, C., Gobo, G., Gubrium, J., and Silverman, D. *Qualitative research practice*, 2004. Sage Publications: Thousand Oaks, CA.

[24] Robson, C. *Real World Research*, 1993. Blackwell Science, Oxford.

[25] Rubin, H., and Rubin, I. *Qualitative interviewing: the art of hearing data*, 1995. Sage Publications: Thousand Oaks, CA.

[26] Seaman, C. *Qualitative Methods in Empirical Studies of Software Engineering*, 1999. *IEEE Transactions on Software Engineering*, 25 (4), 557-572.

[27] Taylor, S., and Bogdan, R. *Introduction to qualitative research methods*, 1984. John Wiley & Sons, NY.

[28] Weiss, R. *Learning from Strangers*, 1994. Free Press, NY.