

Interviews: Qualitative Approach

Introduction to Interview Methods

The purpose of an interview is to ask questions to get someone's point of view. Coolican (2004) identifies five different interviewing styles:

1. **Nondirective.** Nondirective interviews are conversational, and participants talk about anything they wish. The interviews in many therapy sessions are nondirective, where the goal is not so much to gather data for a study as to help participants understand themselves. Results from nondirective interviews are often case studies of therapy sessions.
2. **Informal interviews.** Informal interviews are also conversational in nature, but frequently the researcher wants to gather data for a study. Although there are no set questions or a set response scale, the researcher has a general topic to cover. If the participants get off topic, interviewers use prompts to get them back to the research topic.
3. **Semistructured interviews.** *The semistructured interview is popular with researchers who take a qualitative approach, where the goal is to identify important themes for future research.* Coolican (2004) writes that researchers conducting semistructured interviews have a specific topic to investigate but questions are not asked in the same way to all participants. The goal is to understand the person's way of making meaning in a particular context. This leaves the interviewer free to engage in a conversation with the participant, and the participant might automatically cover the topics the interviewer is investigating. The interviewer will go back and ask anything the participant did not answer spontaneously. Legard, Keegan, and Ward (2003) add that semistructured interviews are meant to be flexible, interactive, probing beyond the surface level (with the researcher asking follow-up

questions for this purpose), and generative (creating new knowledge). Semistructured interviews can be **one-to-one interviews**, where the interviewer talks with one person at a time, or **focus groups**, where a group of people with a particular expertise are interviewed together. Some semistructured interviews make use of **narratives**, a particular type of story that helps people make meaning of their culture. Narratives make up a lot of human conversation and make sense as a research technique. Narrative-style interviews encourage participants to tell their story when answering questions. Sometimes narrative interviews make use of **vignettes**. Sue Arthur and James Nazroo (2003) describe vignettes as "hypothetical examples" (p. 129) that give focus group studies consistency. The vignette represents a typical case that someone might come across. Focus group participants then narrate (tell their story) about the vignette. Vignettes are useful for making sure that everyone in the focus group is familiar with the examples for discussion.

4. **Structured but open-ended interviews.** Sometimes interviewers ask preset questions to standardize interview procedures. However, these interviews still allow participants to respond in any way they wish.
5. **Structured interviews.** Structured interviews collect data with a predetermined scale, such as a **Likert Scale**. Structured interviews have several advantages. They are more objectively verifiable. The researcher needs less training to administer them. **Researcher bias** from emotional involvement is less likely. A disadvantage of structured interviews is that they lack the richness of less structured interview formats. *Researchers who wish to quantify the results of interviews often use this style.* **Email interviews** and **telephone interviews** can use this format. Anyone who is collecting data through email or from an Internet source must consider special ethical issues in addition to those discussed in Chapter 3. For example, how do researchers know if they really have informed consent for online research (Azar, 2000)? Participants check agreement forms, but how do researchers know that participants understand the intent and requirements of the study?

Which type of interview should a researcher select? *It depends on the goals of research.* All have a valid place in psychological research.

Interviews are sometimes designed and evaluated with the purpose of quantification. For example, a researcher might want to correlate responses from interviews to individualism or collectivism, predicting that either individualism or collectivism is related to the way that someone responds.

In contrast, an interviewer might wish to identify themes coming from interviews, where the goal is to understand how participants make meaning in the context of their particular life circumstances. This goal is important for qualitative uses of interviews.

Next are three examples of qualitative interviews. Each study illustrates a different way to design interview studies.

Note to IB Students

These interview studies are also useful for the Paper 2 learning outcome for abnormal psychology about the etiology of mental disorders.

Interview Study 1: One-to-One Semistructured Interview Investigating Attitudes Toward Disordered Eating (Combined with a Quantitative Field Experiment)

Becker, Burwell, Gilman, Herzog, and Hamburg (2002) studied **disordered eating** in Fijian adolescent girls before and after the widespread availability of television. As is sometimes the case, Becker and colleagues used both quantitative and qualitative methods. The research question was whether viewing Western television contributes to eating disorders in non-Western cultures.

Becker and colleagues first ran a field experiment. Female teenagers at two secondary schools in Nadroga, Fiji, filled out questionnaires in 1995, when people had limited access to television, and then again in 1998, after several years of greater access to Western television.

The second part of the study, the qualitative interviews, is more important to our discussion of interviewing. After data collection in 1998, Becker and colleagues selected a **purposive sample** of girls from the group that had filled out the questionnaire (the EAT-26, reviewed in detail in Chapter 9). They collected narrative data from a subset of these girls representing a range of disordered eating attitudes, eating behaviors, and television viewing. Interview questions probed attitudes and behaviors about dieting and

perceptions of weight and body image in relation to Fijian traditional culture. Some example questions were “How do you feel about your weight?” “Do you want to look different from the way your parents think you should look?” “Do you admire any characters on TV?”

Several **themes** emerged from the interview data:

1. The girls admired the television characters and wanted to be like them by changing their hairstyle and body shape.
2. Fully 83% felt that television had influenced the way they and their friends felt about their body shape.
3. Another 40% believed that they could enhance their career prospects if they lost weight, and 30% believed that TV characters were good role models for how to behave at work.
4. All the girls felt that television affected attitudes toward traditional culture and that now there was a growing intergenerational conflict surrounding food consumption. This conflict was especially noticed in girls who had altered their attitudes about food consumption to reflect a Western attitude, with 31% reporting that their parents wanted the girls to eat more food than they wanted to eat.

Interview Study 2: Semistructured Interview about Male Body Image Using Focus Groups

Duane Hargreaves and Marika Tiggermann (2006) investigated **male body image**. Both experiments and questionnaires have been used to investigate body image, but the results are mixed. Most samples use females, and the methodologies are not the most suitable for understanding a boy's unique perspective. **Focus groups** were selected over one-to-one interviews because it was thought that the boys would disclose more about their views in the comfort of a group.

The study had two aims. The first was to describe body image investment, defined as “the degree of cognitive and behavioral importance that people assign to their body and appearance” (p. 568), and body image evaluation, defined as someone's satisfaction or dissatisfaction with his body. The second was to explore factors that were unique to male body image,

particularly the role of the media and whether or not it was acceptable for boys to talk to others about their body image.

Three focus groups were created from 28 boys, aged 14 to 16, from an Australian public school. The focus groups, formed around grade level, included ninth, tenth, and eleventh graders.

A semistructured interview was used because the researchers wanted the boys to discuss the subject with one another in an informal discussion. Interview questions included “Do guys your age care about what they look like?” “How do guys your age generally feel about the way they look?” “Do guys compare their looks to friends/media?” “Do boys ever talk to other people about these sorts of things?”

Five **themes** emerged from the interviews.

1. *Body image investment.* The boys did not generally worry about their appearance unless they were trying to impress girls, and when they did care, they were reluctant to admit it.
2. *Boys' appearance ideal.* The groups said that an ideal male body was buff (muscular and strong) and that the desire to be buff was motivated by a number of things, such as sport competition, impressing girls, and self-defense.
3. *Media and body image.* The groups said that the media were a source of fashion far more than the source of an ideal to judge one's body against. The peer group was identified as a more important influence. But only a small number of boys said they would do anything to become more buff. Instead, they were more likely to change their hair or clothes to fit in with peers.
4. *Was it okay to talk about male body image?* Appearance and body image were not considered acceptable topics. This was partly because it was not judged as important and also because of the fear of appearing feminine.
5. *Body image evaluation.* While most of the boys felt all right about their appearance, there were some participants who expressed concern about their body image, but in specific ways, such as size of muscles. Some boys said that their bodies were still changing, and they thought their bodies were changing more to the ideal (very different from what girls report).

The authors recognized that the focus groups may have kept some boys from reporting some of their greatest concerns about body image, but they felt that the study was important because it identified new themes.

The study's small sample from an Australian public school included little ethnic diversity, so the authors warned readers to be careful about **generalizing** the results outside the sample.

Interview Study 3: Semistructured Narrative Interviews Investigating Lay Perceptions of Depression Using Focus Groups

Elialilia Okello and Seggane Musisi (2006) studied lay perceptions of psychotic **depression** in the Baganda of Uganda and how cultural practices shaped views about appropriate treatments. The authors wrote, “Culture is the lens or template used in constructing, defining, and interpreting reality” (p. 61). Depression is one of the most prevalent mental disorders in Uganda, so it is critical for health care providers to evaluate whether using a Western individual style of health care services is appropriate in Uganda.

Semistructured **narrative interviews** were used with both individuals and the **focus groups**.

Okello and Musisi used **purposive sampling** to select participants for the focus groups, based on age, gender, and ethnocultural background. The authors believed that age and gender might be important factors influencing perceptions, so the focus groups represented different ages and genders. In addition, one cultural group the Baganda, was used, to “avoid diluting the responses due to ethnic diversity” (p. 63). Participants were selected for individual interviews on the basis of their standing in the community as people who were “opinion leaders” (p. 63). Five groups for individual interviews were created: village elders, traditional healers, local leaders, faith healers, and community health workers (laypersons trained to deliver basic services).

A case **vignette** of psychotic depression (though not given a specific diagnostic label) was read to participants in both the focus groups and the individual interviews. Psychotic depression was selected for this study because depression is common in Uganda, it is often undiagnosed and left untreated, and people are most likely to seek treatment only for the most severe cases.

The vignette described a person with the following characteristics. He or she suffers from a lack of interest in pleasurable activities, is withdrawn, and does not clean or take care of himself or herself. The person claims that the ancestors are unhappy and that the ancestors call for the person to die, calling the person worthless and useless, and blaming the person for past sins. The person sometimes talks to himself or herself or stares. Sometimes the person sees dead ancestors or dreams of them.

All the responses were recorded and **verbatim transcription** was used. Here are some of the results.

1. After participants heard the vignette, the discussion started with “What do you call this condition” (p. 65)? Participants in the focus groups and individual interviews gave similar answers, with most saying the symptoms were related to eByekika (clan issues/problems) and Lubaale (ancestral gods). When the interviewers *probed* the response, participants said that the terms were interchangeable and that the symptoms should be treated according to cultural practices.
2. Questions about the cause of the disorder generated the most discussion about why someone had mental disorders, including where someone might get help. Five categories were created:
 - a. Neglecting traditional values
 - b. Breaking taboos
 - c. Mixing foreign and traditional beliefs
 - d. Abandoning traditional beliefs and religion
 - e. “Lost blood,” such as burying a family member outside traditional burial grounds, causing the haunting of the family member
3. Questions about seeking help generated some interesting discussion about treatment. First, medical and/or traditional healers were named as the best source for help with physical symptoms. But a complete cure was believed to come from persons with access to the supernatural realm. Traditional healers and diviners are the only people with access to the supernatural realm. Participants were firm that hospitalization had no role to play in treating this illness.

The authors concluded that the study had important implications for treating depression in the Bagandans, including appreciating the lay understanding of mental disorder and understanding the need for social support from the family/clan system in recovery. A Western individual style of delivering services would be greatly underutilized by this cultural group.

Evaluate Semistructured, Focus Group, and Narrative Interviews

Advantages of using semistructured interviews include the following (Coolican, 2004):

1. A natural conversation provides a rich account of a person’s situation.
2. Questions can be adapted to probe the context and meaning according to the natural flow of the conversation.

3. Participants are more likely to give detailed responses if they are relaxed.

Disadvantages of using semistructured interviews include the following (Coolican, 2004):

1. The **reliability** of semistructured interviews can be poor. It is hard to compare data from one study with data from the next, as each one uses different questions and probes. In addition, interviews take place outside a controlled setting, so it is hard to replicate the situations. There is no one way that qualitative research should be conducted, so design differences make it challenging to compare studies.
2. Interviews usually use small samples that are often limited to one culture or one school because of the time it takes to conduct them.
3. It takes more training to conduct an unstructured interview.

Advantages of focus groups include the following (Neuman, 2006):

1. People are more likely to express their real opinions in a natural setting.
2. Participants hear others’ responses and might be reminded of something important. The group setting can help participants clarify their own opinions.
3. People feel empowered, as if their opinions count.
4. I add to Neuman’s list the fact that interactions among focus group members can give the study more richness and depth.

Disadvantages of focus groups include the following (Neuman, 2006):

1. An individual’s attitudes may become more extreme after participation in a focus group; this result is called the **polarization effect**.
2. Usually only one topic can be discussed at a time in focus groups. In questionnaires or surveys, many topics can be addressed at once.
3. Individuals have less time to talk, so individual opinions are not necessarily developed.
4. The focus group moderator may unknowingly limit what the group discusses (**researcher bias**).

Advantages of narrative interviews include the following:

1. Narratives allow people to tell their story. Narratives make up a lot of human communication, and they are the *primary way people make meaning* of their culture. Storytelling is a rich form of interview data. Okello and Musisi (2006) saw more meaning expressed when participants were allowed to develop stories through vignettes.

2. Narratives provide a window into someone's life context.

There is one disadvantage of narrative interviews: narratives are very personal, and the individual stories from an entire group may present coding challenges to the researcher.

Discuss Considerations Involved Before, During, and After an Interview

Legard and colleagues (2003) and Coolican (2004) suggest that researchers consider the following questions before an interview.

1. What qualities and training are important for an interviewer? Interviewers should be good listeners, logical thinkers, curious, interested, respectful of others, and they should have good memories. The best interviewer has good rapport with participants and "displays a sense of tranquility—an inner stillness which communicates interest and attention and which is accompanied by a feeling of being comfortable with the interviewee and the situation" (Legard et al., 2003, p. 143). The more unstructured the interview, the more important thorough training becomes.
2. The interviewer should learn as much as he or she can about the language and culture of participants.
3. How should a researcher direct the interview?
 - a. Stage 1 is the arrival, where the researcher must make sure the interviewee feels comfortable and in control of the setting.
 - b. Stage 2 involves introducing the interview with clear statements about the purpose of the study and getting the consent of the participant to continue.
 - c. Stage 3 is the beginning; it involves gathering information that provides context, such as relevant background information.
 - d. Stage 4 is the working part of the interview, where the participant is led through the key topics identified for the study. Here the participant works at a deeper level, such as remembering things not considered for a long time.
 - e. Stage 5 is the ending, where the discussion returns to the surface level.

- f. Stage 6 occurs after the interview. The tape recorder is turned off and the interview ends with thanks to participants for their help. Perhaps participants need reassurance that their participation is confidential. If a participant remembers something important after the formal recording is over, perhaps he or she will need to say it again into the tape recorder.
4. Ethical issues are considered. Participants give consent, are guaranteed confidentiality, and are allowed to change their opinion at any time or even to end their participation.
5. Good interview questions and probes are developed. For example, interviewers should ask clear questions, avoid double-barreled questions (where two things are asked at the same time), and avoid leading questions. There are many different probes appropriate for interviews. One example of a probe is amplifying a previous comment from a participant, such as "Can you tell me a little more about . . . ?"
6. The potential for **demand characteristics** (where cues from a researcher make the interviewee anticipate and respond to what he or she thinks the researcher wants) and **researcher bias to influence the study** are considered. For example, an interviewer's gender, personal qualities, or ethnicity might increase the chances of demand characteristics coming into play.
7. Details of scheduling interviews are considered. How long will the interview last? Where should it be held? Is the setting for the interview safe for the participant and the researcher? How will data be recorded? Should anyone else be allowed to attend?

During an interview, Legard and colleagues (2003) advise researchers to consider the following points:

1. The interviewer must create and maintain good relationships with participants.
2. There are two things to consider about recording the interview. If the researcher stops taking notes, is this sending a message to the participant that his or her comments are unimportant? If video-recording equipment is used, to what extent is it dominating the room?
3. The interviewer must continue to think about the staging process.
4. Throughout the interview, the researcher must be aware of the fine line between showing empathy and becoming overly involved with participants. For example, if a participant expresses a position or feeling that is different from that of the researcher, it is important that the researcher not express his or her views and feelings. An em-

pathetic position is best, which involves probing to find out more about the participant's position and feelings.

5. The researcher may need to manage sensitive issues in two areas. First, the topic may be sensitive. The researcher has consent to discuss the topic but may need to lessen a participant's embarrassment or unease with empathy, such as saying, "I know this is difficult, but could you tell me . . . ?" Second, a participant may show a strong emotional response, such as anxiety or anger. The researcher can ease anxiety by acknowledging that the topic is difficult and sharing some general information about how other people have felt about the topic. A researcher should not take it personally if a participant becomes upset or angry; instead, the researcher should defuse the situation by acknowledging the feelings and allowing exploration.

After an interview, Legard and colleagues (2003) advise researchers to consider the following points:

1. The researcher must ensure that the end of the interview goes smoothly. Participants should feel that their help is appreciated, should have information about finding out the results of the study, and should have information about how to withdraw data if desired.
2. Researchers must decide how to transcribe data and then code it into thematic categories. There are two choices, **verbatim transcription** (traditional transcription), where the exact words are transcribed and nothing else, and **postmodern transcription**. In verbatim transcription, researchers transcribe the exact spoken words of subjects. It is quicker than the postmodern approach. Verbatim transcription does not take into account subtle and often important nonverbal communications, such as voice tone, posture, and pauses. In postmodern transcription, researchers record words and pauses, along with the "um's." This gives richness to traditional transcription. The researcher must decide before the study which of the subtle nonverbal messages to record, such as voice tone, posture, and pauses. The entire interview is transcribed, not just the surface words. There are some potential weaknesses to the postmodern approach. It is desirable to have high interrater reliability between the transcriptions of independent raters. Interviews must also be transcribed soon after data are collected so that researchers do not forget the subtleties observed in the interviews. The discussion under the next learning outcome provides details about analyzing interview data.

Explain How Researchers Use Inductive Content Analysis (Thematic Analysis) on Interview Transcripts

All three of the preceding examples of semistructured interviews used **inductive content analysis** to analyze data. Induction is a "bottom-up" approach whereby important **themes** emerge from data analysis of interview transcripts. Although there is no one accepted way to analyze data from qualitative studies, over time guidelines have become more precise (Neuman, 2006).

Qualitative data analysis has the following goals (Neuman, 2006). These goals are very different from data analysis in quantitative studies.

1. Qualitative data analysis is *less standardized* and *more inductive* than in quantitative data analysis.
2. Qualitative researchers think about *emerging patterns* as they gather data and might even probe interviewees about new ideas that emerge. In quantitative research, data analysis takes place only after all the data are gathered.
3. The goal of qualitative analysis is to *create good generalizations* about a person's or group's experience. These generalizations are considered plausible explanations, and researchers give supporting evidence to minimize alternative explanations.
4. Qualitative researchers analyze data by *organizing it into categories according to important themes* that emerge from thinking about the data. All the interpretations are grounded in the data rather than in a theory that might be guiding an experimenter.
5. While quantitative data is organized into numerical units for statistical analysis, qualitative researchers *organize raw data from transcripts into categories and themes that make the data manageable*. **Coding** is central to organizing raw data into coherent categories based on themes. Coding is hard work and is time-consuming. On the next page, you will find a four-stage process for coding. To sum up, the researcher wants to make meaning out of the transcripts and uses a coding system to create categories about important emerging themes.

One popular method for inductive content analysis is **interpretative phenomenological analysis (IPA)** (Willig, 2001). IPA first surfaced as a philosophical concept but became popular with psychologists because it gave them a way to understand a person's unique experience of the world. *IPA is a researcher's interpretation of someone's experiences of a phenomenon.*

Researchers using IPA make some assumptions about the nature of knowledge. IPA is concerned with a person's subjective reality. IPA assumes that individuals can have different subjective experiences about the same objective reality and that these subjective experiences are possible because all reality is filtered through each person's unique beliefs and expectations. The researcher's role is to understand the subjective experiences, and to do this, the researcher must become engaged with the participant's experience (meaning that the researcher is not objective).

IPA involves a series of four stages for how a researcher creates meaningful categories from the themes that emerge from semistructured interview transcripts.

1. In stage 1, researchers read and reread the transcripts. This is a time of reflection. The researcher takes time to comment and ask questions about the interview transcript. Notes at this stage are unfocused; their purpose is to document first impressions.
2. In stage 2, researchers identify and name specific **themes** that emerge from the transcript. Psychological terms are used to give general labels to the themes, such as "loss" or "self-efficacy."
3. In stage 3, researchers impose structure on the general themes from stage 2. Sometimes several general themes come together naturally into categories (clusters of themes). If this happens, the researcher comes up with a name for the category, such as "cultural expectations," "media influences," or "childhood experiences." Some of the categories might be grouped together, as, for example, "the early years."
4. In stage 4, researchers create a summary table of the theme categories along with supporting quotations with a notation on where the material is located in the transcript. A summary table for a study on attitudes about disordered eating might look like this.

Cluster 1: Cultural Expectation		
Traditional culture	"what we ate at the festival"	lines 6 and 7
Parental vs. child beliefs	"what parents thought appropriate to eat"	line 9
Television models	"the thin girls get more glamorous jobs"	lines 10–12

IPA is not just for interviews. For example, Ma (2008) used inductive content analysis in her case study on eating disorders. Ma's study is detailed in Chapter 8.

Observations: Quantitative and Qualitative Approaches

Introduction to Observation Research

Observation is used in both experimental and nonexperimental research, so it can be used quantitatively or qualitatively. Gillham (2008) identifies many styles of observation, including the following:

1. **Structured observation**
2. **Semistructured observation**
3. The use of **observation techniques in experiments**
4. **Unstructured observation**, such as **ethnography**
5. **Self-observation**, such as **diaries**

Quantitative Use of Observation

Researchers sometimes want to quantify the results of observation studies. For example, it is useful to know the frequency of children's aggressive behaviors in play, which can be correlated to gender or other variables. At other times, experimenters measure a dependent variable through observations. For example, Bandura's (1965, 1973) Bobo experiments measured aggression after exposure to different conditions with an observation checklist. But do not call Bandura's Bobo studies "observation studies." The Bobo studies are lab experiments. Data were just collected with observation checklists.

Structured observation grids are developed to quantify observations. Items on the grid cover all possible behaviors related to the research

question. An example research question is “To what extent is the level of aggression in play different between males and females?” Aggression in play is then made **operational**, meaning made concrete and observable, as shown in the following grid.

Aggressive Play in Males and Females		
Aggressive play	Male	Female
Yells		
Punches		
Kicks		
Threatens with toy gun		

An actual grid is lengthier in order to record all behaviors related to aggression. Every instance of the observed behaviors gets a check on the grid.

Researchers decide ahead of time whether to use **time**, **event**, or **point sampling** (Coolican, 2004; Goodwin, 1998). Researchers using time recording do not try to keep a continuous record of everything. Instead, behavior is sampled at predefined times. Event sampling means recording a specific set of events and ignoring all others. Preserving the details of high school graduation events is an example. Point sampling means observing an individual’s behavior before moving on to the next person in the sample.

It is hoped that **interobserver reliability** (or interrater reliability) is high—that more than one researcher sees the same things. Interrater reliability reduces **researcher bias**. Bandura’s Bobo experiments used more than one observer, and there was high reliability between the raters. Reliability is calculated as a correlation coefficient, and coefficients closest to 1.0 are the strongest.

Qualitative Use of Observation Research

Three different qualitative studies are detailed next. Each illustrates a different way to conduct observations.

Note to IB Students

The first study is relevant for the Paper 2 health psychology option about prevention strategies and treatments for substance abuse and addictive behavior. The second is useful for the Paper 2 abnormal psychology option about examining the concepts of normality and abnormality. The third study is helpful for the Paper 1 learning outcome about evaluating two models or theories of one cognitive process.

Observation Study 1: Overt Participant-Observation about Addiction Recovery (Combined with Semistructured Interviews)

Joanne Ehrmin (2002) studied the recovery of substance-dependent African-American women through observation and semistructured interviews. The aims of the study were to explore the recovery care needs of the women living in an inner-city transitional home for substance abusers and help them successfully complete treatment.

Ehrmin believes that there are three aspects of these women’s recovery that require research methods focused on context. First, it is well documented that substance-abusing women are likely to have experienced abuse and rape before they start using the substances. Second, many women say they use drugs and alcohol to numb the emotional pain of their experiences. Last, little research has explored other factors that contribute to these women’s substance abuse, particularly the death of their mothers when the women were young and racism. Quantitative research would not help our understanding of these experiences.

In addition, different cultural groups have unique recovery needs. This view comes from the “culture care diversity and universality theory of nursing.” This theory suggests that care is best provided within the context of the patient’s beliefs and values. It challenges a Eurocentric view of care in favor of one that is more appropriate for African-American patients.

Ehrmin took on the role of **participant-observer** in her **ethnography**, meaning an observation study of cultural practices (in this study, the cultural context of African-American women). The study used **overt observation**, meaning that participants knew they were being observed.

Ehrmin spent three years studying a transitional home in a large U.S. midwestern city. She attended the Friday afternoon “house meeting” and then gradually lengthened her stay to include supper. Eventually she worked her way into participating in many of the woman’s activities, such as attending AA meetings, cooking meals, having meals with the women and their children, and celebrating birthdays.

Ehrmin called her level of participation that of a “**moderate participant observer**,” meaning that she kept a balance between participation as an insider and participation as an outsider. In addition to taking observation notes, Ehrmin kept a **reflexive journal** to document her own feelings and biases.

Ehrmin used an **opportunity sample** of 12 key and 18 general participants for the participant-observation portion of the study. The key participants were the most knowledgeable. The rest included staff and other women recovering from substance abuse who lived in the community but had not stayed at the transitional home.

Three to five interviews were then conducted with 11 of the 12 key participants. This part of the study used a **purposive sample** of the most knowledgeable women from the key participants. All interviews were audiotaped and used **verbatim transcription**.

Guaranteeing **confidentiality** was of particular importance because of the potential legal consequences from the substance use. Ehrmin assigned participants numbers and used the numbers in the coding phase of data analysis to prevent anyone from being identified.

Ehrmin’s study has a high level of **credibility**. Here are three things that Ehrmin did to increase the credibility in her study.

1. Ehrmin increased the certainty of the findings by observing participants over a three-year period and documenting the important context issues throughout the study.
2. Ehrmin increased the ability to **generalize** findings with a purposive sample that met specific guidelines and also gathered rich detailed data (over 1,000 pages of transcripts just for the interviews).
3. Ehrmin kept a **reflexive journal**.

The results included two main **themes** and numerous subthemes, some of which are listed here.

1. One main theme was that the women needed to work through the emotional pain of their abuses and losses.
2. A second main theme was that working through emotional pain gave the women an understanding of the context for their use of drugs and alcohol.
3. One subtheme was the death of loved ones.

4. A second subtheme was racism.
5. A third subtheme was rejection.

Ehrmin concluded that working through emotional pain was a key factor for the recovery of African-American substance abusers because of their high rates of using drugs and alcohol to escape from difficult life situations. The women needed treatment that helped them experience pain rather than numb it. In addition, Ehrmin determined that these women faced “cultural pain” associated with racism and discrimination that should be part of understanding the context of their recovery.

Observation Study 2: Covert Participant-Observation Research about Mental Health Diagnosis

David Rosenhan (1973) conducted a **covert observation** to investigate problems in diagnosing and classifying mental disorders. Although this study is older and Rosenhan does not use the language of modern qualitative research, it is an excellent example of covert observation and very popular with students. Rosenhan’s study is free on the Internet.

Rosenhan asked the question “Do the salient characteristics that lead to diagnosis reside in the patients themselves or in the environments and contexts in which observers find them” (p. 1)?

Eight people were recruited to attempt admission into 12 hospitals in five states on the East and West coasts. The pseudopatients were three psychologists, a pediatrician, a psychiatrist, a painter, and a housewife. Three were women and five were men. All adopted pseudonyms. Would the pseudopatients get admitted? How would they be treated? How long would it take to get released?

All pseudopatients called the hospitals and made appointments. Upon arrival, everyone gave the same symptom—that he or she heard voices that said “thud,” “empty,” and “hollow.” Other than this fake symptom, pseudonyms, and false employment information, all other information was real, such as details about family relationships and a variety of typical life frustrations and successes. Everyone was admitted, and all but one person received the diagnosis of schizophrenia. The length of stays varied from 7 to 52 days, with an average of 19.

Once admitted, none of the pseudopatients showed any symptoms of mental disorder. So no one's insanity was detected because of abnormal behavior. In fact, many of the regular patients voiced suspicion that the pseudopatients were fakes.

The staff made regular notes about the pseudopatients, and an example of how they were viewed follows. One pseudopatient described his normal relationship with his parents to a staff doctor. He described a close relationship with his mother but a more remote relationship with his father. However, as a teenager, the man and his father became close. He was presently married and had a warm relationship with his wife, with only a few disagreements over the years. Rosenhan thinks there was nothing pathological about this description, but when the doctor wrote the case history of the pseudopatient, the tone suggested mental illness. For example, the doctor wrote that the man had "a long history of considerable ambivalence in close relationships, which began in early childhood. A warm relationship with his mother cools during his adolescence. A distant relationship with his father is described as becoming very intense" (p. 5).

Rosenhan also found that when real patients went "berserk," the staff treated it as a symptom of their illness, rather than inquiring into any situations that might have contributed to the outburst, such as mistreatment from staff. In addition, there were strict lines separating staff and patients. Staff left their glassed-off stations, which they referred to as "time out of the cage," only 11.3% of the time. The "time out of the cage" was rarely spent interacting with patients. Rather, it was spent fulfilling staff duties.

When pseudopatients addressed staff with typical questions, such as "Could you tell me when I will be eligible for grounds privileges?" they were answered with responses such as "Good morning, Dave, how are you?"

As you might imagine, Rosenhan's study was a serious challenge to the hospitals in two main ways (Hock, 2005). First, it did not appear that people with mental disorders could be distinguished from people without them; classification systems did not help medical practitioners tell who had a disorder and who did not. Second, Rosenhan pointed out that a diagnostic label had severe consequences for the person. The staff treated each patient as if the label represented the real person, ignoring any situational factor that might be relevant to interpreting behavior.

Rosenhan replicated his findings several times between 1973 and 1975. When staff at one hospital claimed that they could not have made such mistakes, Rosenhan said that he would run the study again, but this time a little differently. Rosenhan warned staff that he would send pseudopatients to them over a three-month period. The staff was asked to rate the likelihood that each admitted patient was a pseudopatient using a scale of 1–10. After three months and 193 patients admitted, 41 were judged as fakes. In reality, Rosenhan sent no pseudopatients!

This study is still ethically contentious. A critical ethical concern was the undisclosed researcher roles; hospital staff were deceived. Here are some questions to ask when evaluating the ethics of the study.

1. Did the benefits of the study outweigh the right of the hospital staff to give their informed consent for the study?
2. To what extent was it important to have the trust of those studied?
3. Did Rosenhan protect the confidentiality and anonymity of the staff and hospital properly?
4. Were proper safeguards in place to protect the pseudopatients from harm?

Observation Study 3: Use of Diary-Observation to Investigate Children's Experiences with Stories (Combined with Semistructured Interviews)

In this study about **language**, Alexander, Miller, and Hengst (2001) wanted to understand more about children's experiences with stories. Stories are an important source of information about cultural expectations. As a result, children seem to develop close "emotional attachments to stories and the cultural beliefs and practices which surround such attachments" (p. 1). There is already a large body of research on children's narratives (special stories that aid in understanding cultural practices). But it is just assumed that besides aiding in cultural understanding, the children have emotional attachments to these stories and that these emotional attachments aid children in managing day-to-day stressors and traumatic events. Alexander and colleagues sought more support for the assumption.

Stories are a shared experience between parent and child. The authors expanded previous research about the social significance of children's story attachments by including parent belief systems and practices related to their children and stories. Studying children and parental behavior together is based on Lev Vygotsky's theory that children learn what is expected of them from daily family routines where stories between parents and children make up most of the conversations.

Participants in the first part of the study, the interviews, were 32 families with preschool-age children. They were selected from a participant file at a midwestern U.S. university, perhaps based on who was willing to be in the study and probably an **opportunity sample**. Five mothers from the original 32 families participated in the second part of the study, the diary-observation phase. These mothers were selected because they had expressed interest in the study topic, had the time to record observations in their diaries, and had a good relationship with the researcher. These mothers were paid \$100 per month for two months of participation. Three girls and two boys were observed by their mothers. These participants were Emily (age 3.7), Cheslia (2.7), Isabelle (3.2), Trevor (3.8), and Jeffrey (2.11). The names of the children were changed to ensure **confidentiality**.

Phase 1 of the study consisted of semistructured interviews with the parents and conducted in the home. Researchers audiorecorded the sessions and also took field notes. Researchers told the mothers that they were most interested in stories to which their children showed strong attachments and in which they had a strong interest. The mothers then selected the stories they felt were most important to discuss.

There were three sets of interview questions about these special stories.

1. *Family narrative practices.* Questions asked for information, for example, "How often do you read to the child, what do you read, and in what setting?"
2. *Basic information about the story attachments.* Questions asked included "What are the title and plot?" "How long was the child attached to the story and how did the child express the attachments (such as in his or her play)?"
3. *Mother's beliefs and practices.* Questions included "Why do you think your child is attached to a particular story?" "What were your reactions to your child's repetitive interest in a story?"

The five mothers participating in the diary-observation phase were trained as research assistants. Then they recorded observations of their child's emotional attachments to stories for two months. A diary-observation form was used by all the mothers. The form included the following items:

1. Basic information, such as the title and content of the story, as well as its form (book, video, etc.)
2. A checklist of what the child did in reaction to the story, such as telling/repeating the story to himself or herself or another person, asking questions about the story, listening to and looking intently at the book or video, and creating a new story using the original story characters

3. The feelings the child expressed
4. Behaviors suggesting that the child was tuned in to the story

The **reliability** of the observations was measured to increase the **credibility** of the study. Twice, a researcher and a mother independently made the same observations of the story engagement. The reliability coefficients were high, ranging from .80 to .95 (these are correlations, and the strongest correlations are closest to 1).

Coding of data from both the interviews and the diary-observations was done separately.

Verbatim transcription was used for the interviews. Some details from the interviews were tallied as frequencies, such as how often the parent read to the child. Next, the parts of the transcripts about story attachments were pulled out for more detailed analysis. Eleven **categories** emerged from children's expression of story attachments. They include requesting a story, expressing feelings, sleeping with the book, and pretending/acting out the story.

Important **themes** (the narrative interests of the children) emerged from analyzing the diary-observations. For example, one story reported by the mothers was the video *Land before Time*. Playing with toy dinosaurs was coded as part of the theme "dinosaurs." Another theme involved how the stories helped the children relax. One mother wrote that her daughter could personalize a story about a cat and a dog because they really had a cat and a dog. A last example is the theme of identifying with another's emotions. Several mothers said their children were attached to the story about **Bambi** and cried out every time they reached the part of the story when hunters killed Bambi's mother.

The authors concluded that their findings were similar to the findings of previous case studies on the topic, increasing the **method triangulation** for their themes. The American families had strong attachments to stories, and these stories were part of daily complex social practices the children were learning.

Alexander and colleagues feel that the greatest strength of their study was its **ecological validity**. In addition, several things increased the **trustworthiness** of the interviews and diary-observations. For example, the mothers were trained research assistants who elaborated at length about the story attachments and even admitted to some discrepancies in the data. Even though the mothers considered video watching to be less socially desirable than reading books, they talked equally about their children's use of both, lessening **researcher bias**.

The authors felt that one weakness of the study was that the diary-observations were so time-consuming that the practice limited the number

of mothers who could participate in the study. In addition, the interviews gave researchers less detail about the children's behavior because they relied on the immediately accessible memories of the mothers.

Evaluate Participant, Nonparticipant, Naturalistic, Overt and, Covert Observations

Let's take these observations one at a time.

Taking the role of a participant-observer includes the following advantages:

1. Participant-observation studies have high **ecological validity**.
2. Participant-observation studies reduce problems associated with **participant expectancy**, especially if the observation is covert.
3. Covert participant-observation increases the **reliability** and **validity** of the observations; they are more like real life. •

Taking the role of a participant-observer includes the following disadvantages:

1. **Ethics** is a concern in using **covert observations**.
2. Researchers cannot assume that they blend in with the observed group well enough to eliminate **participant expectancy**.

The role of a nonparticipant observer includes the following disadvantages:

It reduces ethical dilemmas, as **informed consent** can be frequently obtained.

The role of a nonparticipant observer includes the following advantages:

The researcher's presence may increase **participant expectancy effects** and cause participants to alter natural behavior, especially if the study uses **overt observation**.

Conducting naturalistic observations includes the following advantages:

1. **Ecological validity** is high as long as the researcher is pretty clear that his or her presence has not altered natural behavior.

2. Detailed accounts of behavior are possible, as in the diary-observations from the study by Alexander and colleagues (2001).

Conducting naturalistic observations includes the following disadvantages:

1. Ethical dilemmas increase if **covert observation** is used.
2. Data collection can be so time-consuming that only a small number of participants are involved.

Advantages of using covert observations include the following:

Participant expectancy is less of a problem.

Disadvantages of using covert observations include the following:

Ethical dilemmas occur with the **deception** of **covert observation**.

Advantages of using overt observations include the following:

1. **Overt observation** lessens ethical dilemmas.
2. It is the best choice for some types of observations. For example, Ehrmin (2002) needed her role as a researcher to be obvious. She spent a long time getting to know the women so they would feel comfortable during the interviews.

Disadvantages of using overt observations include the following:

Overt observation increases **participant expectancy**.

Discuss Considerations Involved in Setting Up and Carrying Out an Observation

Some basic decisions about the goals of the study and the role of the researcher must be decided first. Gillham (2008) writes that observational studies have several possible goals:

1. Exploratory goals
2. Descriptive goal
3. Evaluation goals

After a goal is selected, Neuman (2006) writes that observation research requires a large amount of preparation.

The following points must be considered in setting up an observation study.

1. Early preparation includes learning about the topic and defocusing. A thorough review of the literature on the topic for study and **defocusing**—emptying the mind of preconceived notions about the research topic—is essential for **credibility**. One way a researcher can defocus is to get a feel for the setting before deciding exactly what to observe. This helps reduce **researcher bias**. In addition, researchers need a lot of self-knowledge before starting the observations. For example, would any personal experience affect one's ability to be open to the experiences in the study? If the researcher had a family member or close friend with an addiction, might this bias him or her to another's experience?
2. The researcher must choose a research site and gain access to it. Three things are important for selecting the best site. First, the researcher wants the field site or group of people that gives the richest data. Second, if conducting ethnography, the researcher wants an unfamiliar site where the researcher has no prior expertise that might influence the **credibility** of the observations. Third, the site must be suitable. The site should be safe and must be physically accessible (e.g., there are no legal barriers to the site).
3. Decide on a researcher role. A researcher's role falls on a continuum between objective and very involved. A researcher might also be a **participant-observer** (blending in with the natural environment either covertly or overtly) or a **nonparticipant-observer** (observer-participant, conducting overt observations and not blending in with the natural environment). For example, Ehrmin (2002) adopted the role of a moderate participant-observer. The researcher must also decide what to disclose about his or her role and purpose. These are important **ethical** decisions.
4. Sometimes researchers need help gaining access to participants. **Gatekeepers**, people with the authority to give the researcher access to a site, can help. Ethnographers may also need **informants**, or people to translate when there are language barriers. Gatekeepers and informants may bias what is observed and interpreted.
5. Gaining access and planning for the actual study involve several steps. Neuman (2006) uses the term **access ladder** to explain the process. Gaining entry to the site is like being on the bottom rung of

the ladder and often involves gathering public data or paying a first visit. For example, at first, Ehrmin attended group meetings for short time periods. Only later did she gain access to the women's private lives. Climbing up to the top rung of the ladder involves establishing and maintaining good relationships with participants.

The following points must be considered in carrying out an observation (Neuman, 2006):

1. Build and maintain rapport with participants, keeping an "attitude of strangeness." Rapport building involves joining in with joyous occasions as well as sharing fears and anxieties. Unless the observation is covert, be aware that some participants will change their behavior during the study; this is called **participant expectancy**. The only way to completely eliminate participant expectancy is to use covert observation. However, sometimes participants get used to the researcher's presence and start behaving normally. But some group members might never be receptive. Neuman calls unwilling participants the **freeze outs**. Not everyone will be won over by the researcher, and it is ethically important to respect this possibility. In addition, researchers sometimes have to manage requests for favors from participants. Researchers must remember that familiarity can blind researchers to the things that are important for the study.
2. Make sure to collect quality data throughout the study. Observing is challenging. The best advice is to collect **thick descriptions**. Thick descriptions are extremely detailed accounts of the setting and context as well as what happened; they are obviously lengthy. A short encounter can take up pages in an observer's notebook. This way, the researcher is less likely to forget what is observed, thus reducing **researcher bias** when the observer recalls events. Otherwise, the constructivist nature of memory may interfere with the "true" observed events. Neuman also reminds observers to embrace **serendipity**. Researchers do not know what is going to happen and should not try to impose their preconceived notions.
3. Take rigorous notes. **Credibility** is increased if others can check and recheck a researcher's notes.
 - a. Take **jotted notes** in the field. It is hard to take good notes in the field, but researchers create a shorthand to remind them of what happened.
 - b. Write **direct observation** notes as soon after as possible (without having talked to anyone about them).

- c. Make **inference notes** when interpretations are made.
 - d. Write **analytic memos**, which are decisions about how to proceed, such as “Make sure to follow up with Jane because she made an insightful comment.”
 - e. Make **maps** and/or **diagrams**, such as a drawing showing where everyone sat at a meeting, so that no one forgets important information.
 - f. Keep a separate **reflexive journal** to make sure that personal biases and feelings are noted.
4. Focus on what is important for the study and sample appropriately. This means that initial observations will direct the rest of the study. In addition, Neuman recommends that observers use **theoretical sampling**.

Researchers have several duties after collecting observations:

1. Conduct interviews with key participants to discuss data interpretations. The study is more **ethical** if participants are allowed to comment on the **credibility** of the recorded data and interpretations.
2. Check and recheck interpretations with other researchers.
3. Debrief participants.

Discuss How Researchers Analyze Data Obtained in Observational Research

Inductive methods are the best choice for analyzing qualitative observation data. **Grounded theory** is one such method. Willig (2001) defines grounded theory as “the progressive identification and integration of categories of meaning from data” (p. 33). The goal of the grounded theory method is to create theory from identifying, refining, and integrating categories and themes.

Numerous strategies are important to grounded theory. Let’s go through them one at a time.

1. **Thick descriptions.** Thick descriptions of observed data are crucial to grounded theory. Only detailed, thick descriptions lend themselves to quality analysis.

2. **Creation of categories.** As with interpretive phenomenological analysis (IPA, detailed in Chapter 6), grounded theory requires that categories be formed around common observations. At the beginning of data analysis, categories are usually at a **low level of abstraction** (descriptive). For example, one low-level category might be “emotions,” including anger, jealousy, or anxiety. Emotions describe what is seen at the observation site. As data analysis continues, **higher levels of abstraction** emerge. These are analytic rather than descriptive. For example, Ehrmin (2002) selected the theme “working through emotional pain,” which reflects a high level of abstraction, to provide a way to interpret the emotions described at the lower level of abstraction.
3. **Constant comparative method.** The researcher moves back and forth between considering similarities and differences among all categories. This way, new categories and subcategories are not left out of the analysis. For example, Willig (2001) suggested that one category that might emerge from observations is “emotions.” It is possible that subcategories will emerge. One example of such a subcategory is emotions that require an object (such as hate). Such a breakdown into subcategories ensures that all types of emotions are identified.
4. **Negative case analysis.** Researchers are always looking for cases that do not fit existing categories. New categories and interpretations might emerge from negative cases.
5. **Theoretical saturation.** Researchers code data into categories and use the constant comparative method until a point is reached when no new themes emerge.