Engineering Psychology

Prof. Naveen Kashyap

Department of Humanities and Social Sciences

Indian Institute of Technology, Guwahati

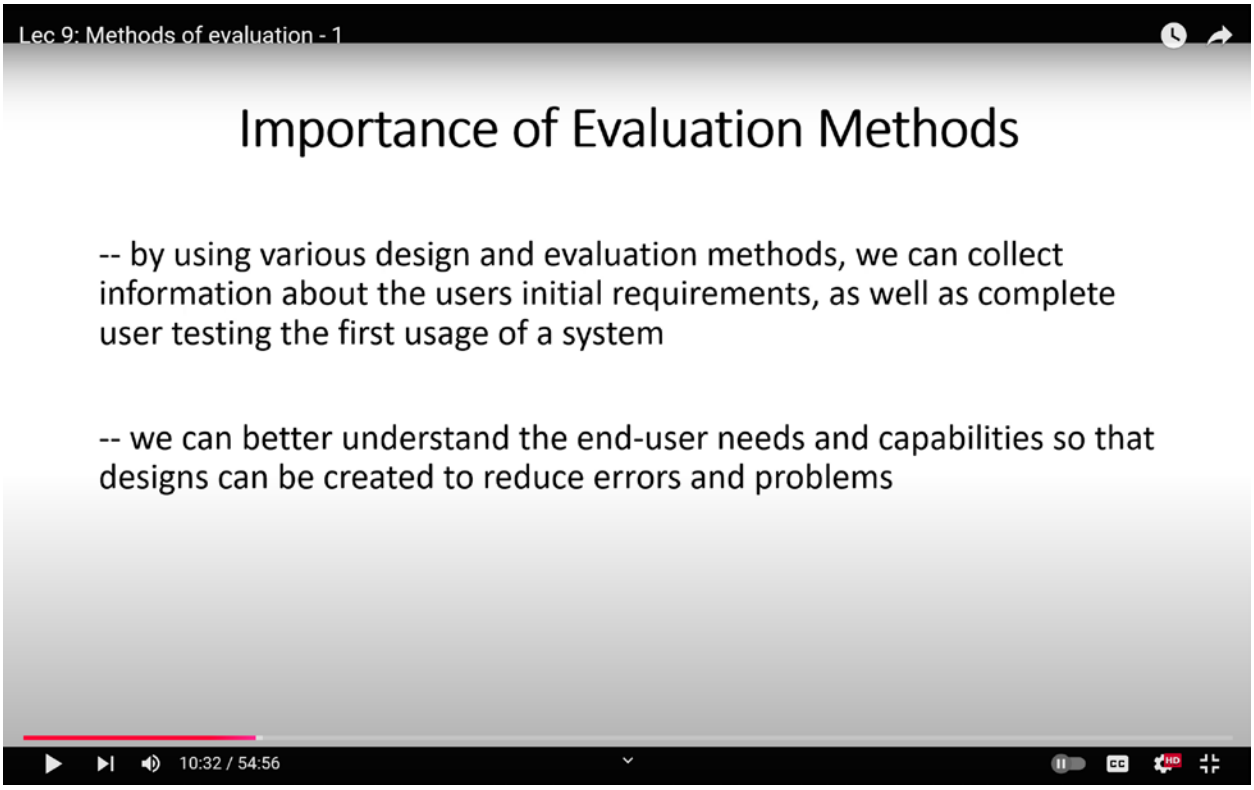
Week-04

Lecture-09

Methods of evaluation - 1

Namaskar, friends. Welcome back to the following lecture in this series for the course on Engineering Psychology. In the previous few lectures, I explained the capabilities and limitations of both the visual and auditory systems. The human visual system and the human auditory system are the two primary sensory modalities humans use to interact with the world around them. I also briefly covered how tactile and olfactory sensory organs can present warnings, alarms, and display information.

(Refer Slide Time: 10:32)



The first two lectures were focused on introducing engineering psychology, discussing its subject matter, and introducing research methodologies that can be used for designing and testing various system-operator interactions.

In today's lecture, and in the following one, I will be discussing the evaluation methods used in engineering psychology. Now that we understand the physiological limitations and capabilities of the human sense organs, how do we evaluate a system designed based on this knowledge? This is the focus of today's lecture. We will explore user-centered design, methods of gathering information from the end users, and the processes for developing these methods to collect valuable insights from system operators and users.

We will examine how to test the usability of a system and determine whether usability principles work with the techniques used for measuring usability. This is a brief summary of the upcoming two lectures.

Let me open today's lecture with a scenario that is common in everyday interactions between people and systems. Most of us book tickets through online websites, and after using them a few times, we become quite familiar with the entire process. We know how to select a ticket, make the payment, and complete the reservation.

Now, assume you are traveling from point A to point B, but for some reason, you cannot find a direct connection. You start thinking about how to make a break journey to reach your destination while keeping the booking cost-effective and the experience smooth. You log in to the familiar website, input your password as usual, and enter the departure and arrival locations. However, since there is no direct flight available, you face a dilemma. How do you determine where to plan your break journey to keep the trip both time- and cost-effective without overcomplicating things?

Nowadays, systems come with advanced intelligence that can suggest several ways to book a break journey. But let's assume you are using a system from an earlier time, where this intelligence isn't available. You haven’t used this break journey feature before, but you know it exists. So, based on your knowledge, you try different combinations to book the ticket, but several factors are in play.

For instance, certain routes are more frequently used, making them cheaper. Some connections are more available than others, and certain time frames offer discounted tickets. Without this

information readily accessible, you rely on your knowledge and the website's assistance to book your ticket, but you end up paying more. Later, you share this experience with a friend who explains how you could have completed the booking with just one click using a hidden feature. You didn't explore this feature because it wasn’t intuitive, and the website didn’t present it upfront. As a result, you spent more time and money on the booking process.

This scenario illustrates a common problem in system design. We frequently see websites and products undergoing upgrades, but these improvements are often so non-intuitive that users struggle to interact with them. From the designer's perspective, these upgrades are meant to benefit the user. However, users often find them difficult to navigate, and these changes can even reduce their performance and effectiveness in completing tasks.

So, what can be done to address this issue? One solution is adopting a user-centered design approach, where end users are involved in every phase of the product development cycle. Engaging end users alongside field experts ensures that the product is designed with the user's perspective in mind, making it more intuitive and user-friendly. This process ultimately helps websites and systems be redesigned in a way that benefits the end user.

In this section, we will delve deeper into why design evaluation is necessary, what user-centered design entails, how data is collected in this process, and how we analyze this data. Additionally, we will explore how usability testing methods are used to assess the effectiveness of these designs.

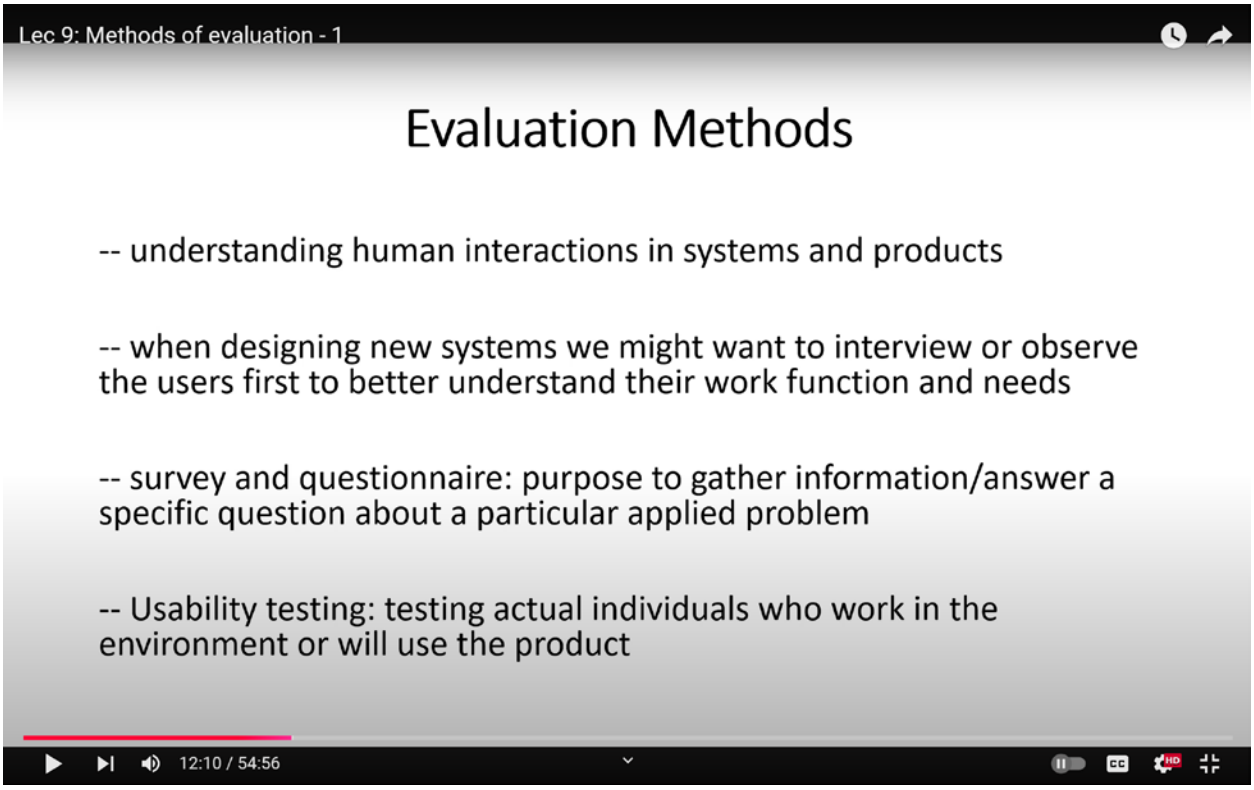
So, let us begin our journey. Now, why is it essential to evaluate methods in engineering psychology? The importance lies in the fact that by using various design and evaluation methods, we can gather information about the user's initial requirements. One primary area for improvement in a design is that previous designs may not have addressed the user's needs effectively. This can also apply to newer products that promise to address gaps left by earlier versions, providing enhanced support to the user.

By employing these evaluative methods, we can ensure that during the initial testing of the product or system, users do not feel lost. When a product or system is launched for the first time, we aim to minimize complaints. To better understand the end user's needs and capabilities, and to design solutions that reduce errors and issues for users, we must evaluate design improvements. The

importance of testing designs lies not only in gathering information but also in understanding users' requirements and assessing how designs can help people commit fewer errors and avoid serious problems.

Evaluation methods in design assist designers in comprehending human interactions with systems and products. Designers are always eager to learn how people interact with systems. While most designs are intended to be intuitive, it is essential to determine whether users understand the mental model or the rationale behind a new design or product. If there is a closer alignment between the intent of the product, the design process, and users' perceptions, the design will improve. Conversely, if there is a disconnect, problems may arise. Thus, one reason for conducting evaluations is to gain insight into how humans interact with systems and products.

(Refer Slide Time: 12:10)



When designing a new system, developers may wish to interview and observe users first to better understand their work functions and needs. All designs, product upgrades, and system modifications are implemented to assist the user. Therefore, it is crucial to identify what the user

requires, how the user functions, what improvements are necessary, and what aspects should remain unchanged. This knowledge is a vital component of the evaluation method.

So, how do we collect data regarding user requirements and understand the work functions and profiles of end users? Two methods are commonly used for data collection: surveys and questionnaires. In essence, surveys and questionnaires serve similar purposes. Surveys can be exploratory in nature, gathering broad information, or they can address specific questions related to particular applied problems.

For instance, if I design a new system that did not previously exist, I might want to know who will use this system and how it will benefit the target users I envision. In such a scenario, the survey would be exploratory. On the other hand, if I discover that a specific process within an existing system confuses users, I might use a targeted questionnaire to identify the source of confusion and determine how to resolve the issue. This approach enhances the intuitiveness of the design and increases the system's usability.

A good example is using the camera on a mobile phone. Users might wonder whether to use portrait mode or general photography mode, what aperture settings to select, or whether to enable the flash. These questions could be addressed by an automatic sensing system. However, information should be provided on how different modes affect the outcome of the same photograph. A brief preview could accompany a specific questionnaire in this context.

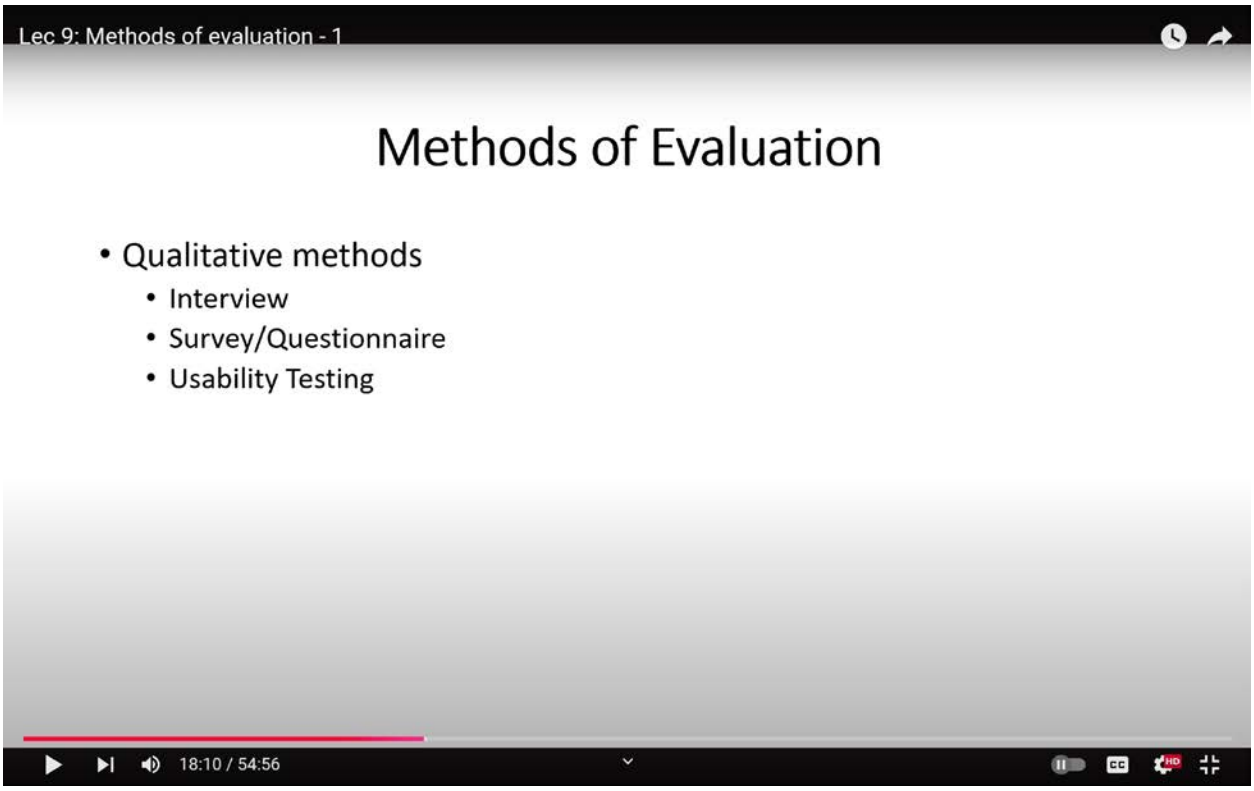
Conversely, if I plan to launch a temperature sensor on a phone but am uncertain whether users will find it useful, I can employ a general survey to gauge initial interest and potential usage frequency. This is important, as significant costs may be incurred in adding this feature.

The next step in evaluation is usability testing. Given that developers and designers have understood how people interact with the system and have grasped the work profiles of users, as well as their needs, data collected through questionnaires or surveys is essential. The next critical phase is to incorporate this information into the new design and conduct usability testing. In usability testing, individuals who operate within the environment are asked to use the product or modifications, and data is gathered while they interact with the product.

By involving users throughout all phases of design development, we can address both known and

unknown issues that may arise with the introduction of new modifications. This approach ultimately saves time and resources for developers and designers. To evaluate and collect data, there are three primary methods: qualitative approaches, such as interviews, questionnaires, surveys, and usability testing. Once data is collected using any of these methods, it is analyzed using various analytical techniques.

(Refer Slide Time: 18:10)

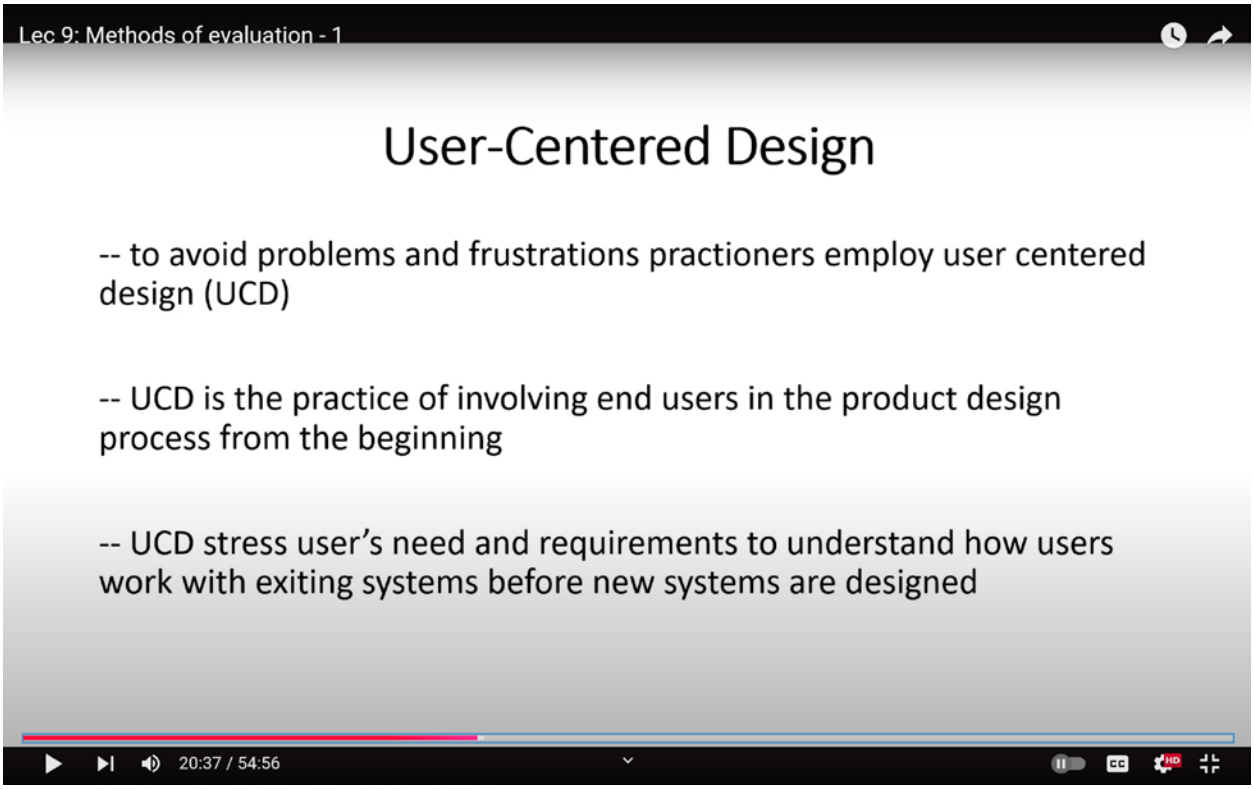


We can utilize descriptive or inferential statistics along with various design approaches to analyze this data. In our previous discussions, we explored the concept of user-centered design, which emphasizes involving actual end users in the development process of a system or product. Let us delve deeper into user-centered design and its significance.

User-centered designs are implemented to prevent the problems and frustrations that practitioners encounter during design iterations. This means that whenever a new iteration of a design is created, designers and developers often experience frustration at various points in the design process due to the need to make assumptions about how the end user will perceive the design and respond to

it. However, by employing user-centered design, where real users are involved and tested at every stage of development, these frustrations can be alleviated. With actual users providing feedback, we can identify potential problems more effectively.

(Refer Slide Time: 20:37)

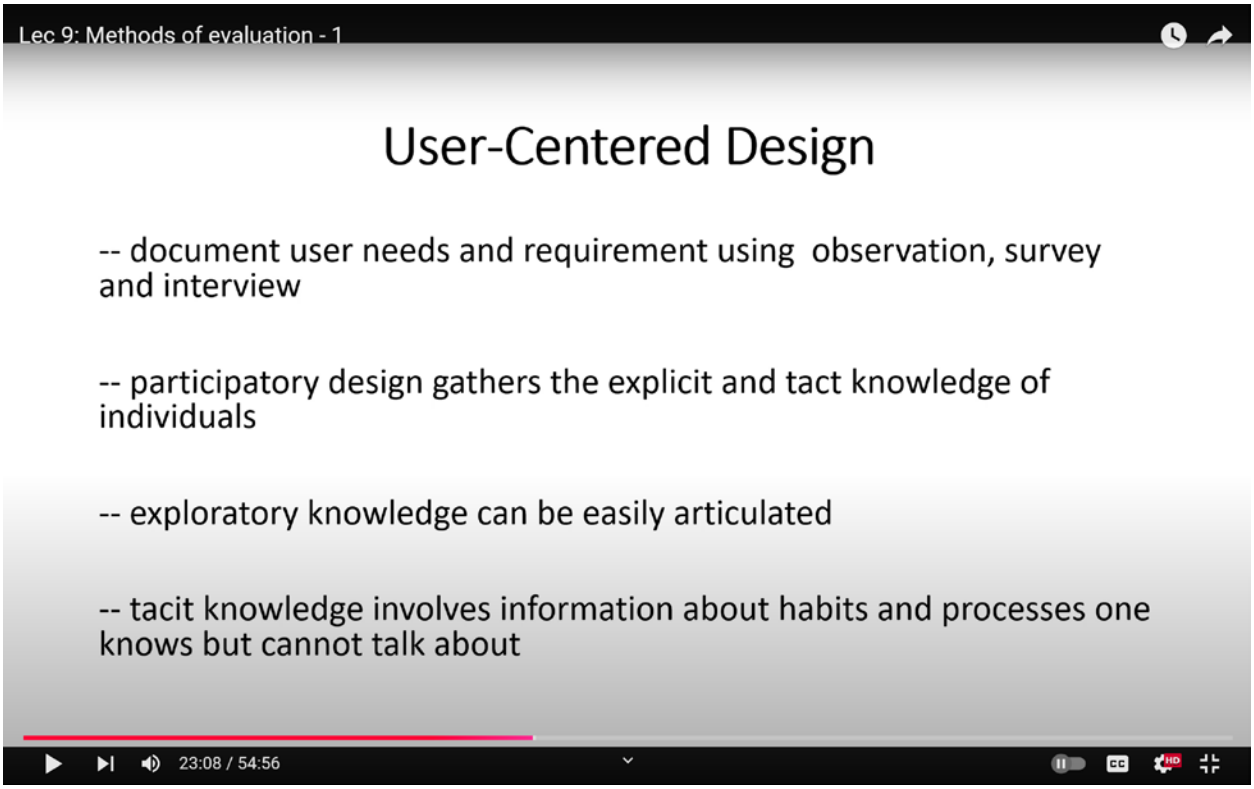


User-centered design involves engaging end users throughout the entire product design process, starting from the initial conception of the design, through the drawing board phase, the development of the mental model of the design, the implementation of prototypes, testing of those prototypes, and culminating in the final launch of the product. This approach prioritizes users' needs and requirements, allowing designers to understand how users interact with existing systems before new systems are developed.

As we discussed previously, user-centered design focuses on examining how individuals interact with older systems, identifying existing problems, and determining what modifications can be made to prevent these issues from recurring. In our first lecture, I explained how Industrial and Organizational (IO) psychology aims to address these problems by training individuals to use and

benefit from the existing system's limitations. In contrast, engineering psychology seeks to modify the system so that users do not need to be trained in areas where they may lack proficiency. Engineering psychology emphasizes the capabilities of users and aims to enhance the interactions between operators and systems.

(Refer Slide Time: 23:08)



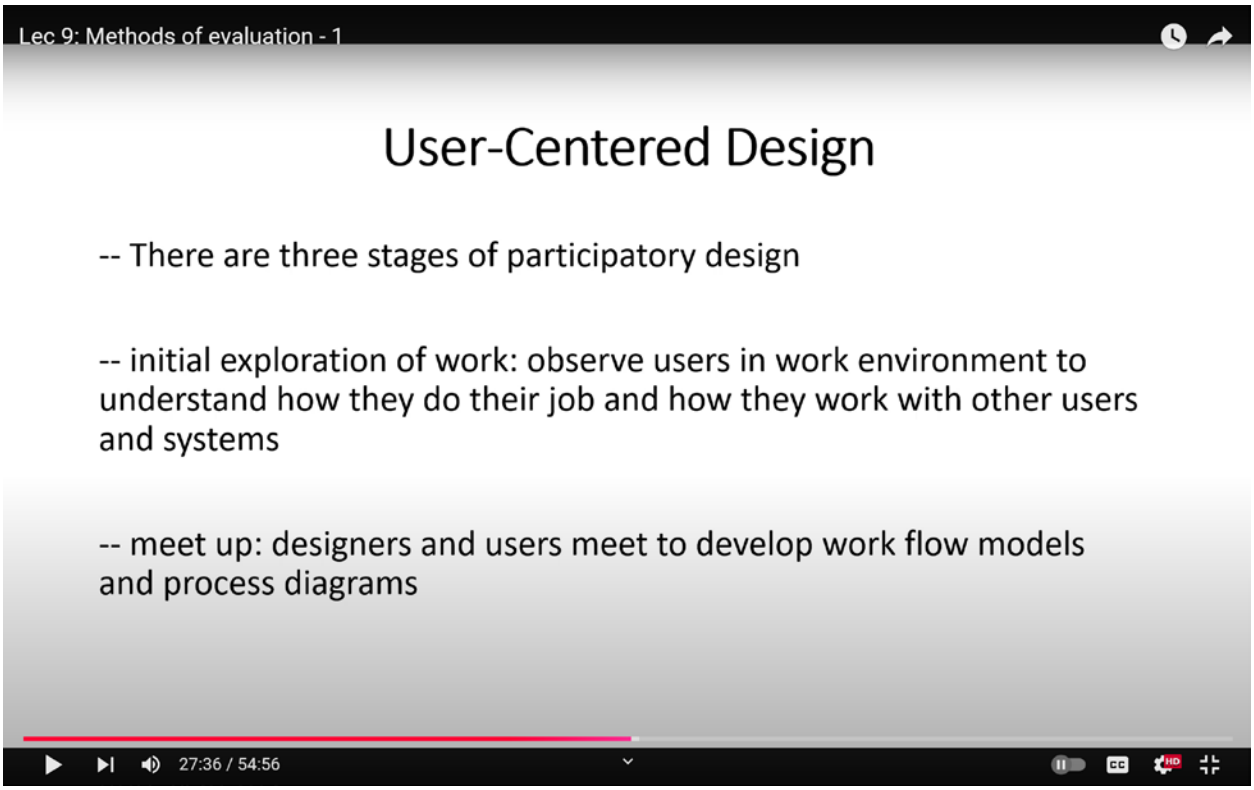
User-centered designs document users' needs and requirements through observation, surveys, and interviews. These are three primary methods for collecting data. In observational studies, subjects are monitored in their natural environment without interference from the observer. Surveys involve asking questions of end users about a system or product, while interviews consist of one-on-one interactions between end users and developers.

One significant advantage of employing user-centered design is the concept of participatory design. This approach allows end users to actively participate in the development and modification of the system or product. Participatory design helps developers gather both explicit and tacit knowledge from individuals. Explicit knowledge refers to what individuals consciously think and

how they interact with the system. This is observable and can be articulated. By observing how users engage with systems, developers can collect information regarding their interactions.

Conversely, tacit knowledge is the implicit knowledge that each individual possesses but cannot easily articulate. A classic example of tacit knowledge is riding a bicycle. While most people can ride a bike, if asked to describe the step-by-step process for teaching someone to ride, they may struggle. Bicycle riding is procedural knowledge rooted in muscle memory and is challenging to explain verbally.

(Refer Slide Time: 27:36)



Tacit knowledge also plays a role in how end users interact with systems, as they may not consciously understand how to describe their actions. For instance, operators of machinery may rely on intuition or mental calculations, but they may not be able to articulate the specific combinations of buttons to press to avoid certain situations. If asked why they made a specific choice, they may not have a clear answer. Instead, their decisions are based on practice, experience, or other forms of non-conscious knowledge.

This kind of information about how users leverage non-conscious knowledge to navigate specific situations can be effectively captured through participatory design, as real users are engaged throughout the process. If designers do not understand how a task was accomplished by an end user, they can ask probing questions to clarify what occurred.

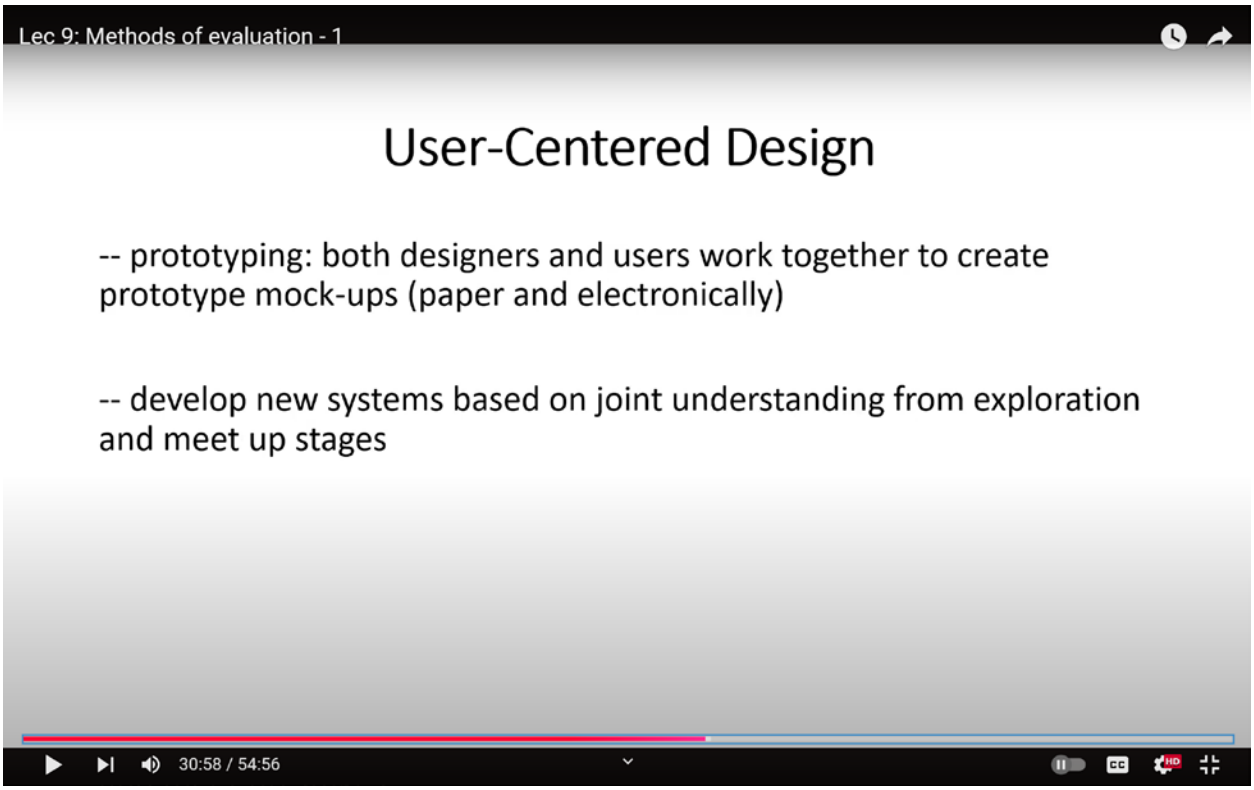
Explicit knowledge can be easily articulated, whereas tacit knowledge involves habits and processes that individuals know but cannot explain. As we have just discussed, this non-conscious form of knowledge is crucial for users in performing specific tasks, yet users may lack awareness of how this knowledge was acquired in the first place.

Most participatory designs consist of three stages. The first stage is the initial exploration of the work. During this phase, designers observe how users interact with the system without interfering with their tasks. The primary goal of this step is to gather extensive knowledge through the observation of users' interactions with the system and to understand their mental models. The mental model aspect will be examined in greater detail during the second stage. However, even in the initial exploration, designers, developers, and experts can gain insights into the users' activities.

In the initial exploration stage, observers watch users in their work environment to comprehend how they perform their tasks and collaborate with other users and systems. This observation includes not only their interaction with their own system but also how cooperative work is conducted. The second step is referred to as the meetup. In this stage, designers meet with users of the system to gain a deeper understanding of the workflow and process diagram that outlines how the system operates. For instance, if I were to redesign a specific section or process in a power plant, the first step would involve visiting the facility to observe how personnel operate within it. This includes tracking the steps they take, from initiating the machines to the final phase of power generation.

After gathering this observational information, designers will sit down with power operators to discuss the entire process of generating power. There is a strong possibility that the observations made will differ from the explanations provided by the operators. Therefore, this second stage is crucial for designers to understand how users interact with systems and the mental models or strategies they employ to accomplish their tasks.

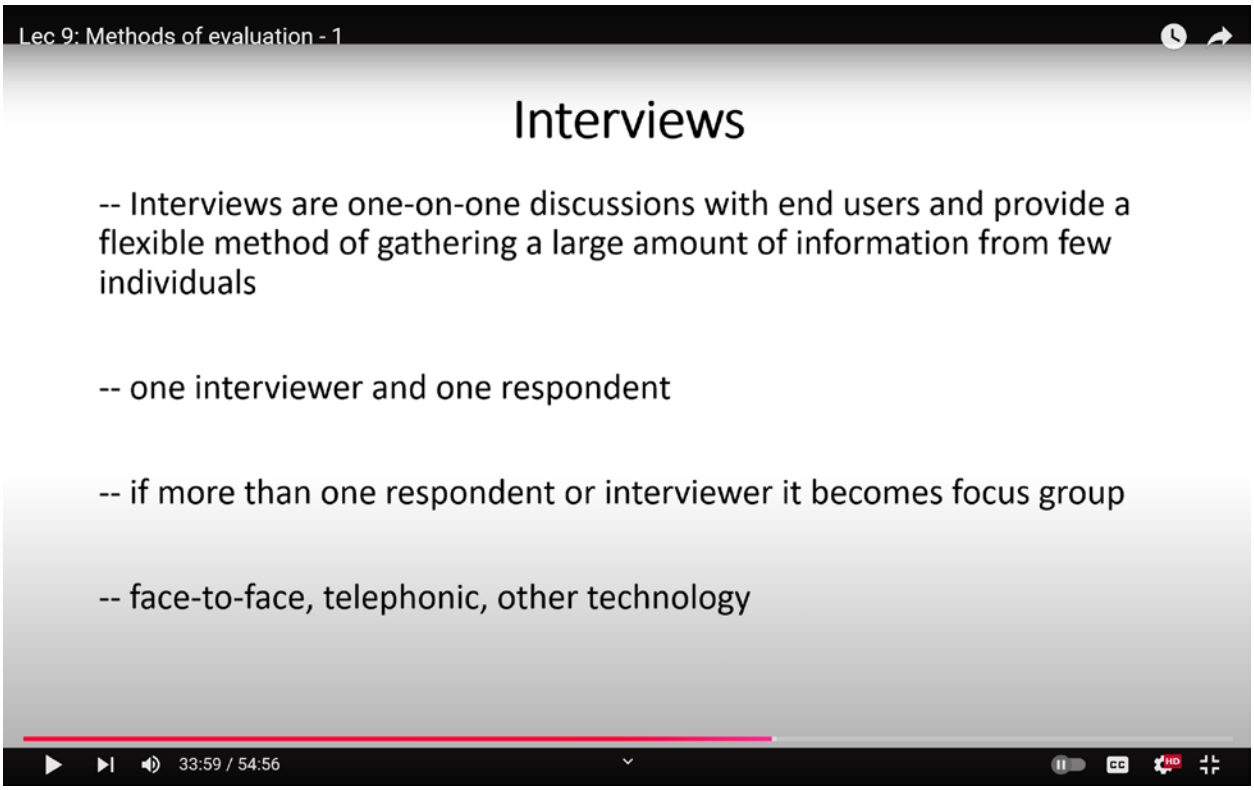
(Refer Slide Time: 30:58)



Once the first two stages are completed, the third stage involves prototyping. In this stage, designers utilize the data collected from their observations, along with the workflow and mental model data, to create new designs or modifications to the existing system. For example, in our power plant case, a modified system design or user interface will be presented to users, who will be asked to interact with it. A mock-up of the new design will be created, which serves as a prototype, not the actual system, but a representation of it. User feedback and performance on this prototype will be observed to determine whether the improvements aimed at enhancing user experience are effective.

In rapid prototyping, both designers and users collaborate to create prototype mock-ups, which may be produced both electronically and on paper. Following this, a new system design is developed based on the collective understanding from the exploration and meetup stages. Thus, prototyping serves as an extension of the observational data gathered in the first step, as well as the interaction data collected during the meetup.

(Refer Slide Time: 33:59)



We have examined how user-centered design operates and how it assists in creating modifications to existing systems, ensuring they are both usable and efficient. Now, we will focus on methods of collecting data, specifically two popular techniques used for gathering information from individuals.

The first method I will discuss is the interview. Most viewers are likely familiar with the concept of an interview, but I will provide a brief overview of what an interview entails, how it can be used for data collection, and its purpose. Interviews typically involve one-on-one discussions with end users, allowing for a flexible approach to gathering substantial amounts of information from a limited number of individuals. Primarily, interviews consist of one respondent and one interviewer, where the interviewer poses questions that the respondent answers.

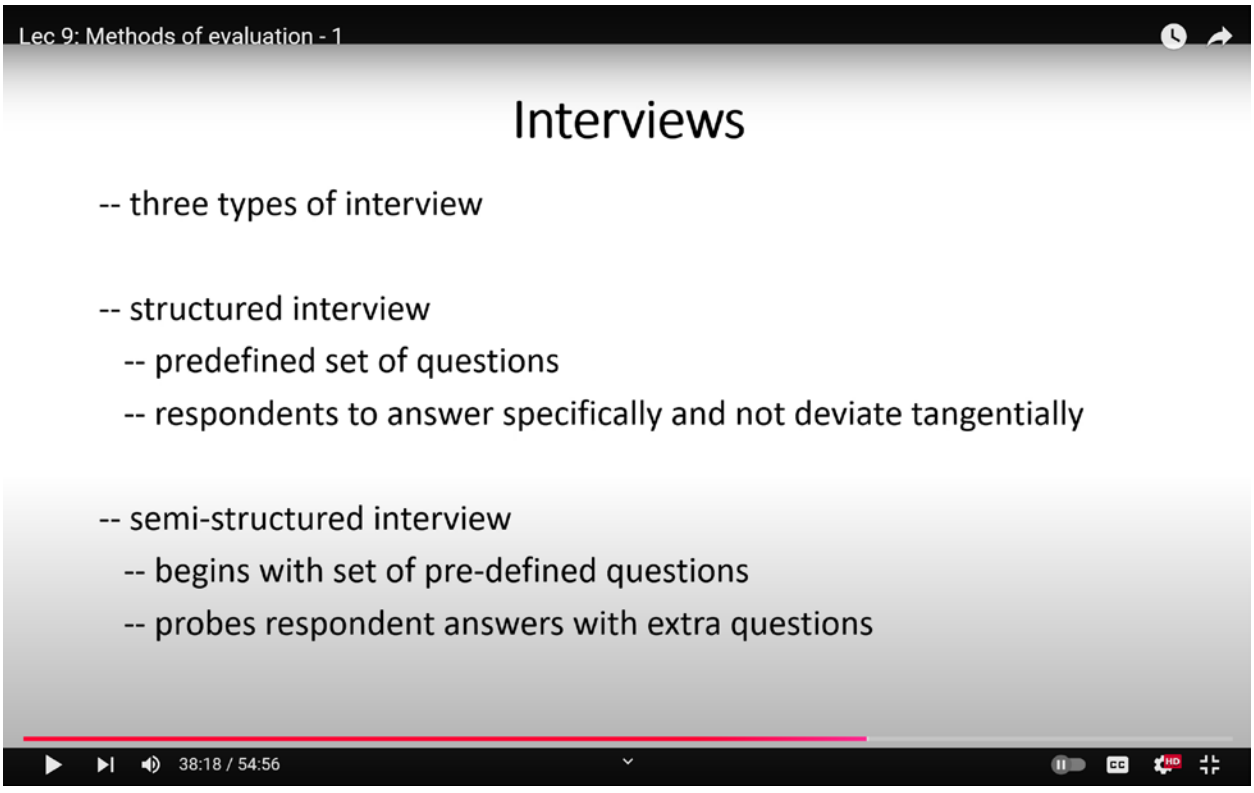
Interviews are commonly encountered in job application processes, where candidates are asked a variety of questions, either in a one-on-one format or in a panel setting. When multiple interviewers are present, it is referred to as a focus group panel discussion or panel interview. However, most

interviews remain a one-on-one interaction.

The flexibility of interviews allows for the collection of both relevant and tangential data related to the questions posed. A considerable volume of information can be obtained from just a few individuals. Typically, in an interview, a question is directed at the respondent, who provides an answer that is recorded. If the interviewer does not fully understand a response, they can ask follow-up or probe questions to clarify and gain a deeper understanding of the respondent's explanation.

An interview generally involves one interviewer and one respondent, as previously noted. If there are multiple respondents or interviewers at any point, the interview transforms into a focus group. The primary distinction between an interview and a focus group is that an interview is characterized by one-to-one interaction, while a focus group involves many-to-one interaction, which can consist of multiple interviewers and a single respondent, one interviewer and several respondents, or multiple interviewers and multiple respondents.

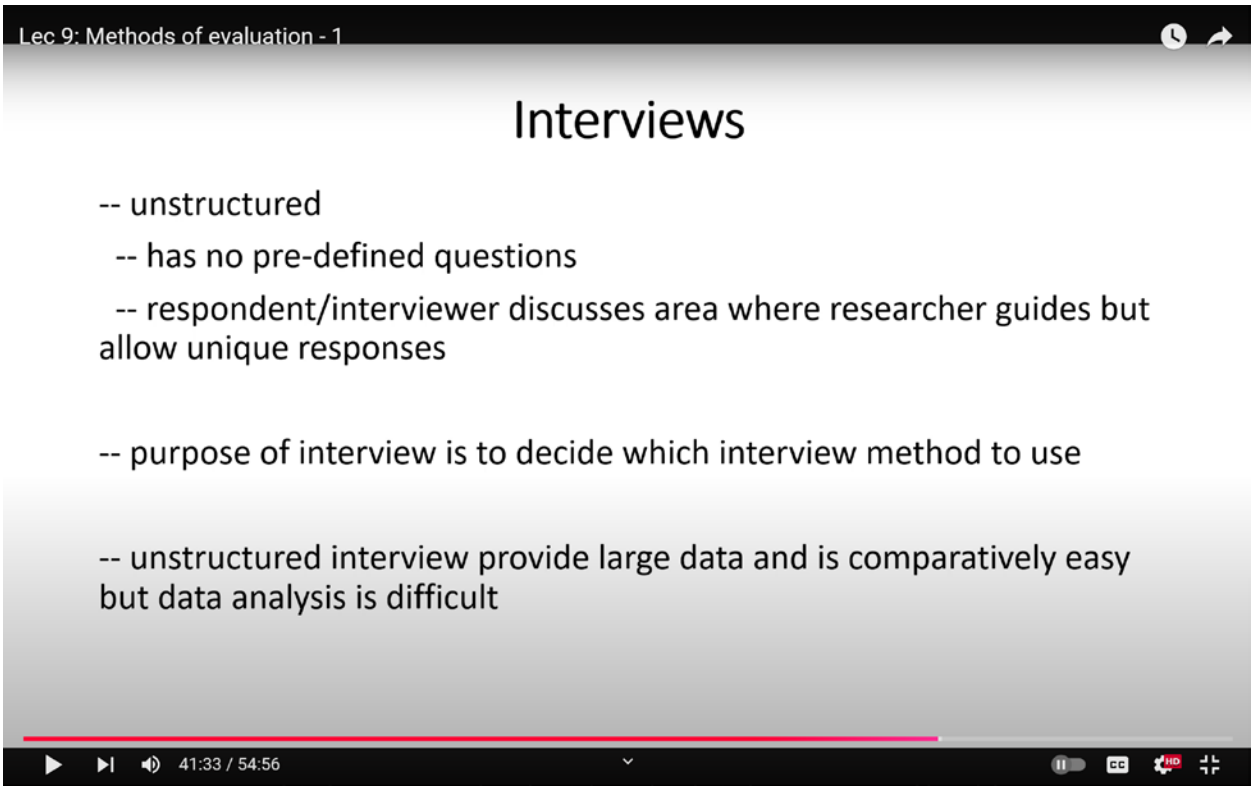
(Refer Slide Time: 38:18)



The number of interviewers and respondents in a focus group should not be large. A focus group consists of a small collection of people brought together for a specific purpose. We will discuss focus groups in detail shortly, but for now, let's explore how an interview is conducted.

An interview can take place face-to-face, where one interviewer and one respondent sit across from each other and engage in conversation. Alternatively, it can be conducted over the phone, wherein the interviewer initiates a call with the respondent, similar to the telephone surveys used by service industries for data collection. Additionally, video interviews are another option. Various methods can be employed for conducting interviews, which generally fall into three distinct types: structured, semi-structured, and unstructured.

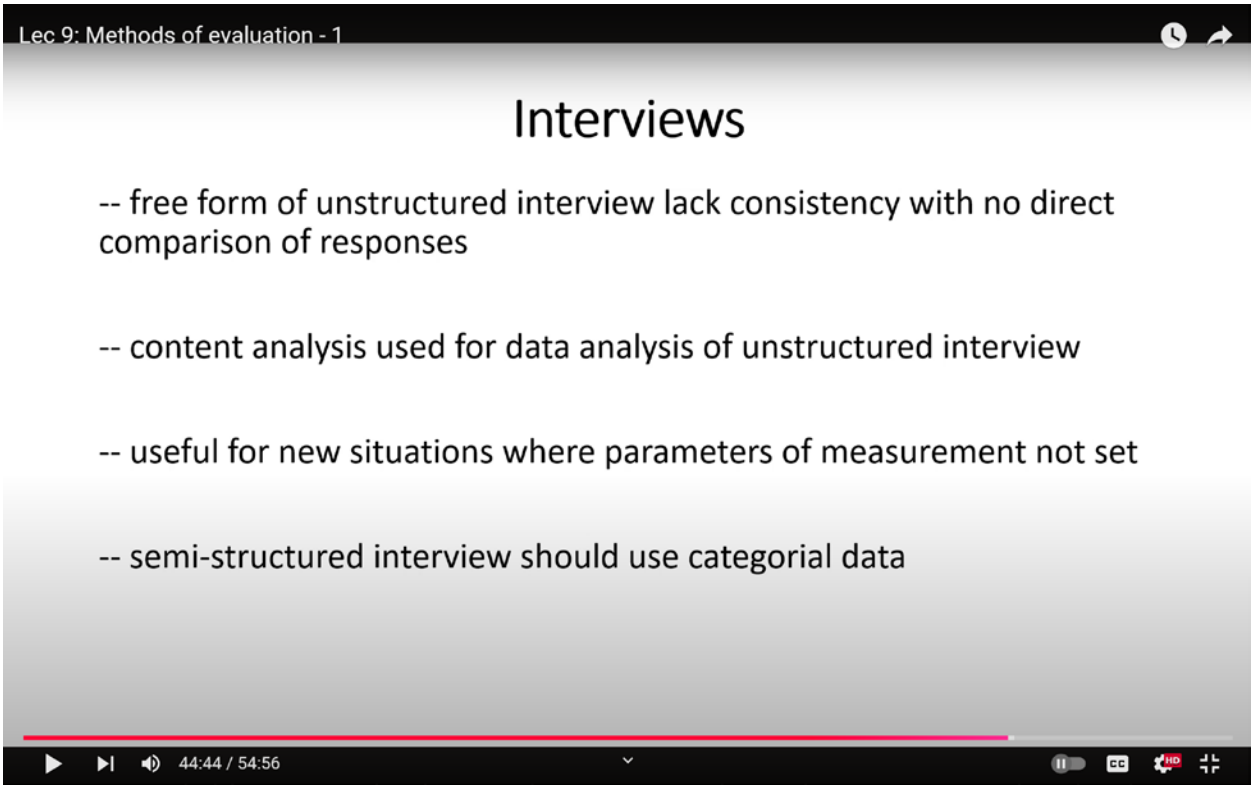
(Refer Slide Time: 41:33)



In a structured interview, a predefined set of questions is presented to the respondents, who must answer these questions. The responses are recorded and analyzed. One critical control in a structured interview is that respondents should answer specifically the questions posed to them and avoid deviating tangentially from the subject matter. For instance, if I ask how to perform a

power cycle in a specific product or system, the answer should pertain directly to the steps required for executing a power cycle, rather than addressing why the power cycle is performed or the necessity of conducting it. Such questions and their corresponding answers will not be entertained; only specific answers to specific questions are allowed.

(Refer Slide Time: 44:44)



In a semi-structured interview, as in the structured format, we begin with pre-specified questions. For example, when defining the power cycle, if the interviewer finds an interesting point while the respondent explains the down phase of a power cycle, the interviewer may decide to probe further. They might ask additional questions regarding why certain actions are taken during the down cycle, particularly if these actions deviate from the manual or standard procedure. The information obtained through this probing can assist in designing better systems.

The third type of interview is the unstructured interview, which lacks predefined questions. This format allows respondents to describe how they work or outline their work profile. In unstructured interviews, individuals can provide a wide range of answers without restrictions or guidance on

what to include or how to structure their responses. The interviewer may guide the conversation but permits unique responses, focusing on areas of interest. Although varied answers are welcomed, excessively vague or unrelated responses may require further probing or guidance from the interviewer to clarify and eliminate ambiguity.

Some control mechanisms are applied in unstructured interviews to ensure the information gathered remains relevant and coherent. Unstructured interviews are considered one of the most effective types because they yield valuable insights often unavailable in structured and semi-structured interviews. The choice of interview method should depend on the desired outcome: if exploratory data is needed, an unstructured interview is ideal; however, if limited and structured data is required, a structured interview should be utilized.

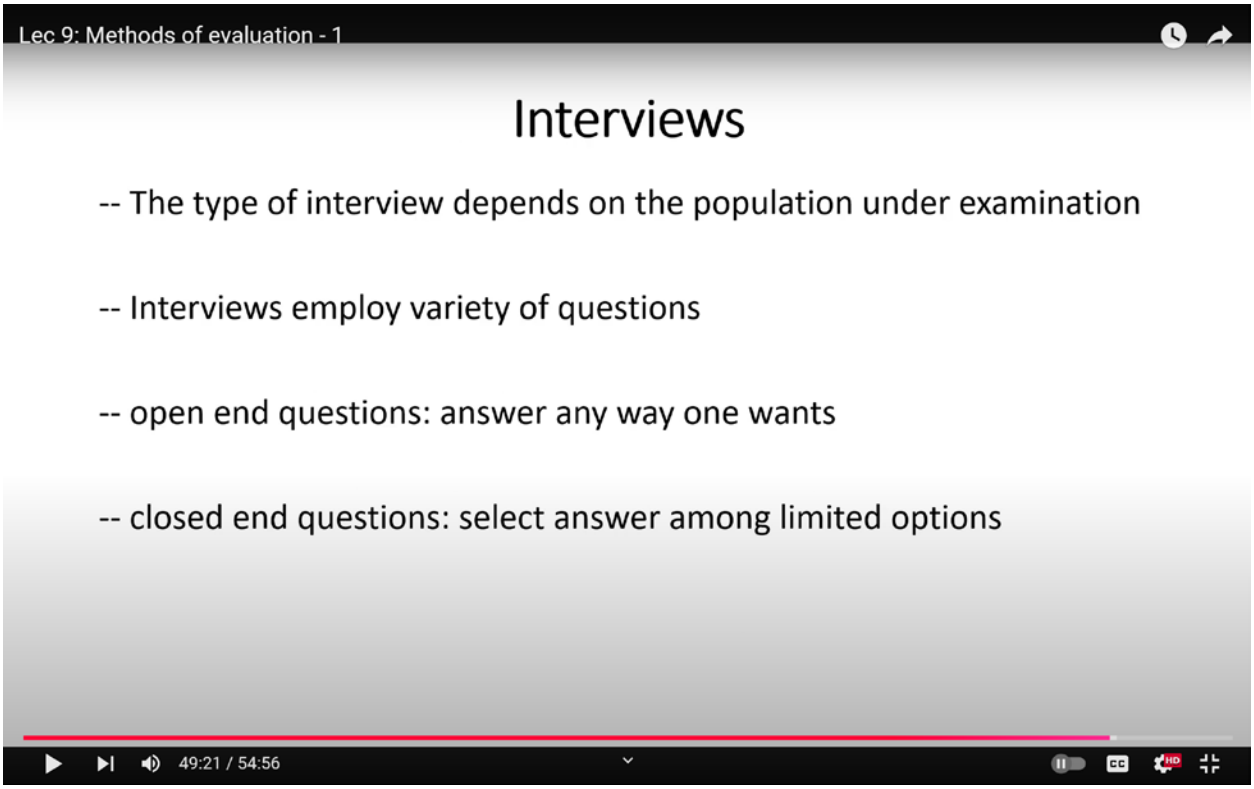
The unstructured interview generates a large amount of data and is relatively easy to conduct; however, data analysis can be challenging. This format allows for extensive input on various topics, and it is easy for participants to respond because the questions are open-ended. Nevertheless, analyzing these answers can be difficult for several reasons. Responses may be vague, out of context, or may contain information related to a non-related phenomenon, which complicates data analysis in an unstructured interview.

Additionally, the free-form nature of unstructured interviews lacks consistency, making direct comparisons of responses problematic. Participants in an unstructured interview can provide a wide range of answers to the same question. For example, when asked to define their happiness index or satisfaction, two individuals may present vastly different measures of satisfaction. One person's concept of satisfaction might revolve around earning money, while another's could be centered on completing their job. Thus, direct comparisons become challenging, as the meaning of satisfaction varies from person to person. The free-form aspect of unstructured interviews means that when asked to define satisfaction, participants’ responses cannot be easily compared across individuals.

One potential method for handling data obtained from unstructured interviews is through content analysis. In content analysis, the raw answers are categorized according to specific themes. A good understanding of content analysis can be developed through the thematic apperception test used in personality assessments. Initially, the responses are collected, and primary measures such as word

frequency, word occurrence, and word usage are analyzed. Each response from participants is then read, and an underlying meaning or theme is generated. A theme refers to what a particular sentence or section of a sentence conveys. Some themes may be pre-established, while others emerge organically from the respondents' answers.

(Refer Slide Time: 49:21)



As the responses are analyzed, it may become evident that different individuals agree on certain aspects, which serves as the foundation for content analysis. For example, if someone states, "My satisfaction depends on how much work I do, the number of hours I come in, and the number of machines I handle," the underlying theme here could be identified as work satisfaction. Such themes are categorized, allowing for comparisons to be made.

This method of analyzing the content of statements provided during an interview is termed content analysis. Unstructured interviews are particularly beneficial in situations where new circumstances arise, and parameters have not yet been established. For example, when a new system or product is introduced and its effectiveness is uncertain, unstructured interviews can serve as exploratory

tools. During these interactions, users may provide developers with insights that reveal potential reasons for utilizing the product, as well as additional thoughts that could aid in marketing efforts. In such cases, unstructured interviews allow for the extraction of dependent variables from the participants' responses.

In semi-structured interviews, categorical data should be utilized, as it is the most suitable type of data for this interview format. Since the questions in semi-structured interviews are predetermined, respondents can first answer these questions within different categories. After a category has been selected, a probing question can be posed to understand the reasons behind choosing that particular category.

Categorical data is best suited for structured interviews. One way to determine which interview technique to use depends on the population being studied and the specific issues at hand. If the target population is large and the issues affect this broader group, then unstructured interviews are appropriate. However, if the population is fractured or stratified, a semi-structured interview may be the better choice. Thus, the nature of the population being addressed and the types of problems they face will define the interview method employed.

Most interviews incorporate a variety of questions, often using different formats. Two common formats of questions used in these interviews are open-ended questions and closed-ended questions. Open-ended questions allow respondents to answer freely, as they see fit. For example, questions like, "How good do you feel after using this new product?" or "How do you evaluate this new modification of the system?" exemplify open-ended questions. Respondents can provide any answer they choose, elaborating on specific features while potentially omitting others. This type of information can help interviewers understand the thought processes of those who developed the modification and whether it is clearly understood by the end users.

In contrast, closed-ended questions offer a limited set of predetermined answers. For instance, if I am making improvements to a part of a system, I might replace a multiple-choice question with a yes-no question. In this scenario, vague answers are not desirable; I require specific responses that directly relate to whether the system modification or test adjustment, such as using a dropdown instead of a multiple-choice option, is more effective.

To illustrate, if Google wishes to assess whether the five-point star rating system previously used for video ratings was effective compared to the current thumbs up/thumbs down system on YouTube, it could utilize closed-ended questions. An example could be, "How do you perceive this change? Is the transition from a five-star system to thumbs up/thumbs down easier for you to understand or more intuitive?" Thus, when examining specific issues, closed-ended questionnaires tend to be more effective than open-ended ones.

Additionally, there are other methods for collecting data, including focus groups and questionnaires, as well as discussions regarding usability testing. We will cover these topics in the next lecture. I look forward to seeing you then. Namaskar, and thank you from the MOOC studio.