Design and Implementation of Human-Computer Interfaces
Dr. Samit Bhattacharya
Department of Computer Science and Engineering
Indian Institute of Technology – Guwahati

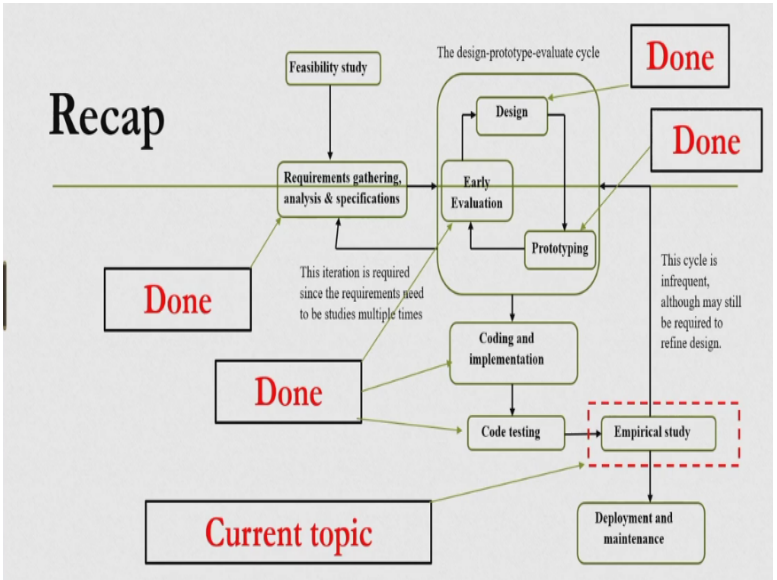
Lecture – 39
Experiment Design - 1

Hello and welcome to the NPTEL MOOCS course on design and implementation of human-computer interfaces. We are going to start lecture number 33 where we are going to start our discussion on another stage of the empirical study that is identification of variables; this comes under experiment design phase. So as is customary before every lecture, we start with a quick recap of what we have learned and where we stand at present in the context of the overall course objective.

Now, in this course, we are going to learn about two things design and implementation of human-computer interfaces. Human-computer interfaces, we have equated with a special class of software systems that are called interactive systems. So, essentially we are interested in two things, design and implementation of interactive systems. In order to achieve these two objectives, we need a systematic approach.

For that purpose, we can follow a development lifecycle. We are currently discussing as part of this course one such interactive system development lifecycle comprising of several stages.

(Refer Slide Time: 02:09)

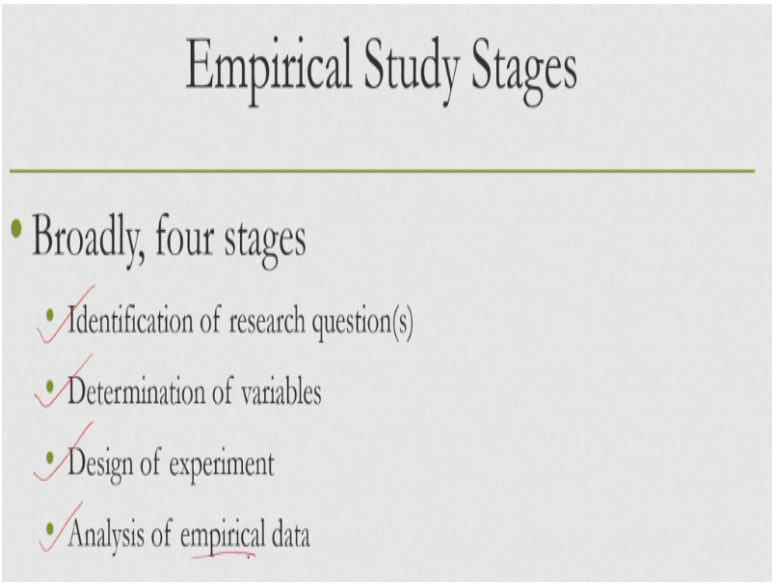


In this lifecycle, we have so far covered requirement gathering requirement gathering analysis and specification stage. Outcome of this stage is software requirements specification document. Then, we have covered the design stage for interface and interaction design, outcome of this is a design document. This is followed by a prototyping phase where we create prototypes for the design ideas that we have come up with.

Then, we go for weak evaluation of our prototype, so that we know whether there are any usability issues. Now, this design prototype evaluate lifecycle continues till we arrive at a stable interface and interaction design. This is followed by design of the code where we create a code design document comprising of modules and the interfaces between modules. We can use either DFT or UML languages to create the document.

Once the code is designed, we go for coding an implementation that is the next stage. So, their outcome is the code of the system. This is followed by code testing where we test for bugs that may be present in our code. At the end we produce a test report. The next stage is empirical study where we study the usability of the end product. Currently, we are discussing the empirical study stage.

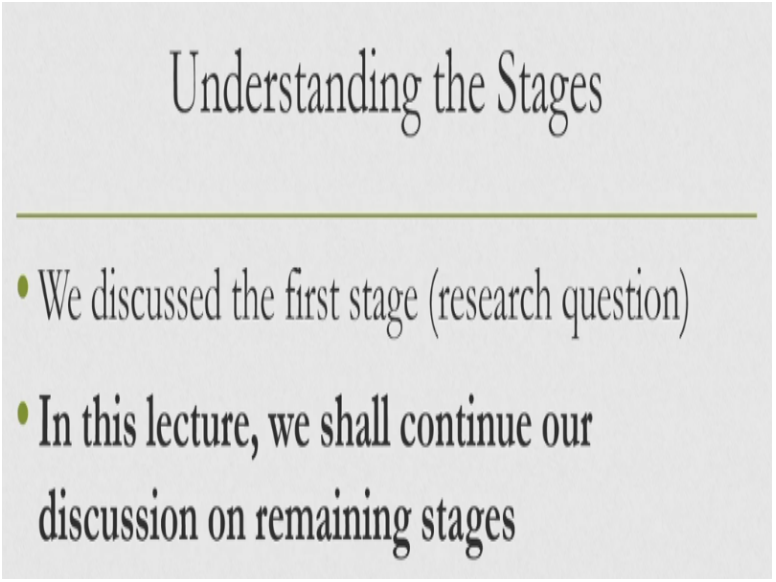
(Refer Slide Time: 03:52)



We have already covered several lectures on empirical study. So, the basic ideas have already been discussed. As we have seen the study is a systematic way of ascertaining usability of a product. Now, it comprises of four stages, what are those stages? Identification of research question that is stage number 1, this is followed by determination of variables stage number

2, this is followed by design of experiment stage number 3 and finally we come to the fourth and final stage that is analysis of observed data or empirical data.

(Refer Slide Time: 04:35)

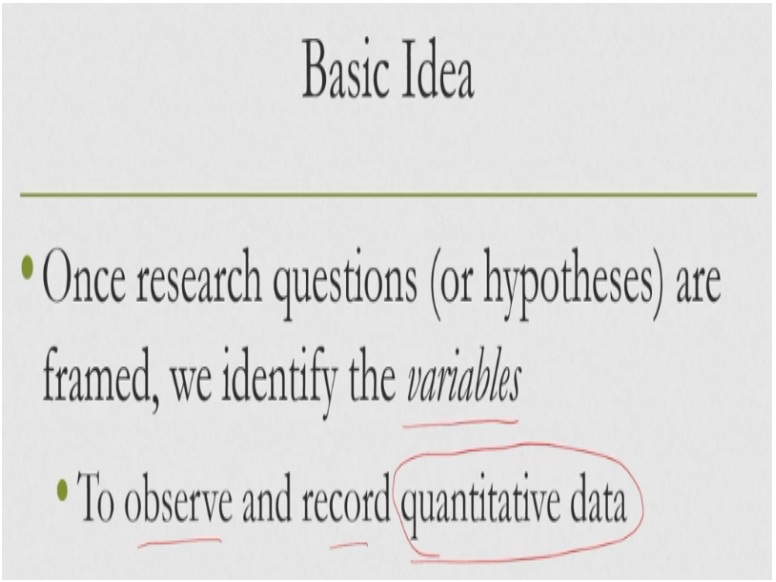


Earlier in the previous lectures, we have discussed the idea of the first stage that is identification of a research question or framing of suitable research questions. Just, to quickly recap why that is important. So, empirical study is all about performing a controlled experiment, observe user behaviour in that experiment and come up with conclusion. Now, the conclusion should be a reliable conclusion which can only be possible if the data that we collect is reliable data.

To collect reliable data, we have to design, we have to plan a very systematic and scientific experiment process where we will observe the behaviour and record the behaviour. In order to know what to observe and what to record, we need to come up with appropriate research questions. Earlier we have seen what happens if the research question is not appropriate and how we can come up with an appropriate research question.

In this lecture, what we are going to do is we shall continue our discussion on the stages of the empirical study and we will start our discussion on the next stage that is identification of variables.

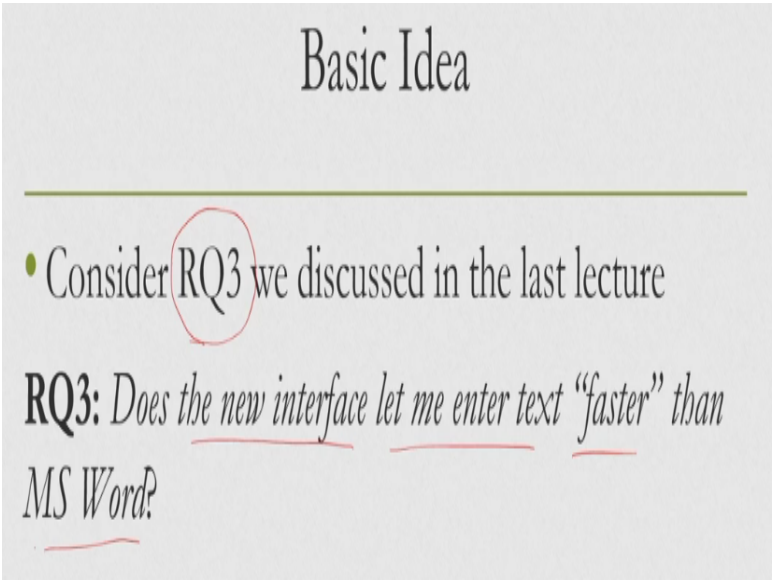
(Refer Slide Time: 06:06)



So, once the research questions are framed, we identify the variables. Now, just to quickly recap, at the end of the previous lecture, we talked about the idea of hypothesis. What we discuss there is that research question is one way of looking at things. Generally, what is done is based on the research questions that we frame, we come up with hypotheses and then try to refute or justify the hypothesis. For each question, there can be two hypotheses.

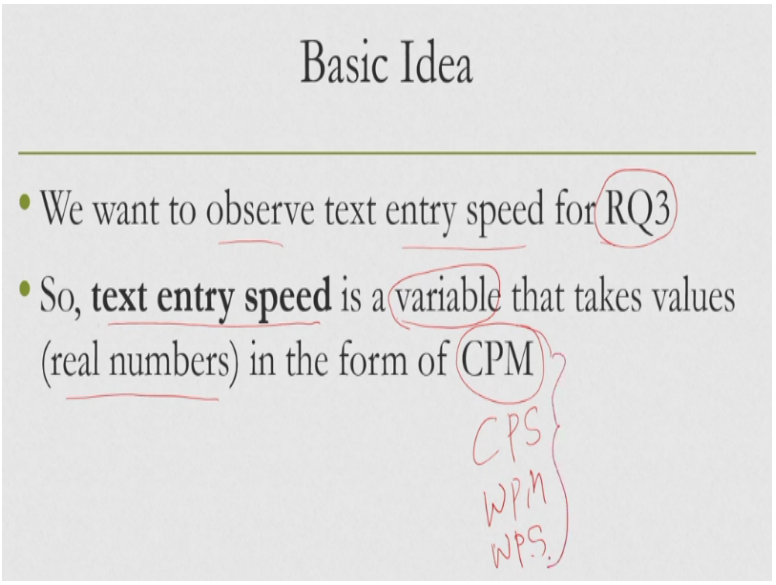
One is the null hypothesis which you wish to refute or nullify, the other one is alternative hypothesis which we aim to establish. So, once the research question or the corresponding hypotheses are framed, we need to identify the variables. Why the variables are needed to be identified to observe and record one quantitative data. So, you want to record some data quantitatively so that we can perform data analysis techniques on them to come to a reliable conclusion about the outcome of the study.

(Refer Slide Time: 07:23)



In order to understand this problem, this issue, let us reconsider the earlier research question that we have framed which we termed as research question 3, RQ3. What was that question? Does the new interface let me enter text faster than MS Word that was the question we frame and we seek answer to this question in the empirical study.

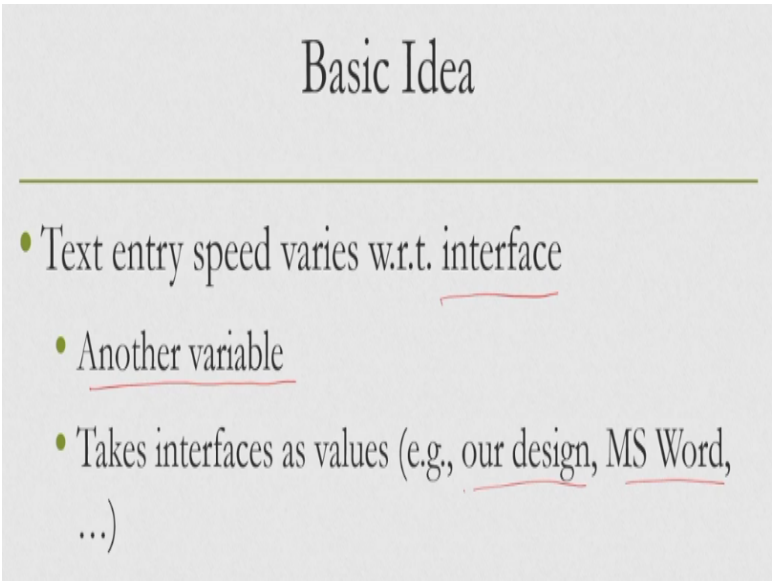
(Refer Slide Time: 07:54)



So, with respect to this question, what we need to do? We want to observe text entry speed for this particular research question to answer this research question. So, here obviously text entry speed is a variable and this variable can take values which are real numbers in this case in the form of CPM or character per minutes, it can be CPS also that is characters per second, WPM that is words per minute or WPS words per second.

So, any of these four variations we can use to record text entry speed. The basic idea is that text entry speed is a variable that can take different real numbers as values and these real numbers signify these things, how many characters were entered per minute or how many characters were entered per second or how many words were entered per minute or how many words were entered per second.

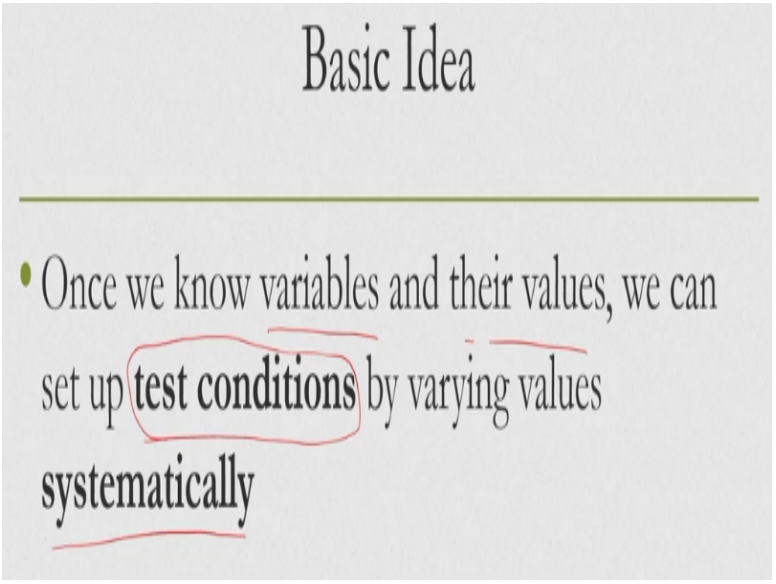
(Refer Slide Time: 09:04)



Another thing we can note here. Text entry speed varies with respect to interface. So, interface can be considered to be another variable. So, we can take the interface as another variable related to this research question and this particular variable the interface variable can take as values different instances of interfaces, for example, our design can be one value, MS Word can be another value.

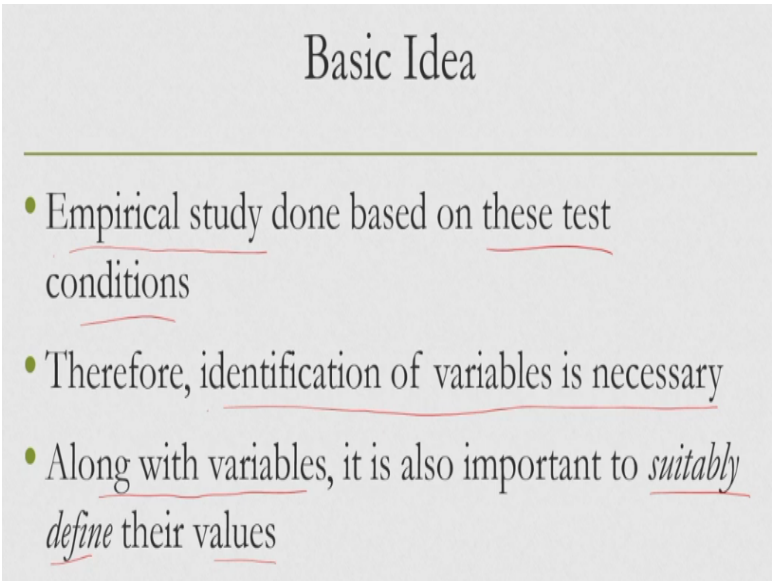
Word Pad can be another value, any other text editor can be another value for this variable. So, you have then two variables identified for this research question, one is text entry speed, other one is interface.

(Refer Slide Time: 09:56)



Once we know the variables and their values, what we can do is we can set up test conditions by varying the values systematically that is a crucial thing that we should remember that the overall idea of identifying variables is to be able to set up test conditions that is assigning values in a systematic manner to these variables.

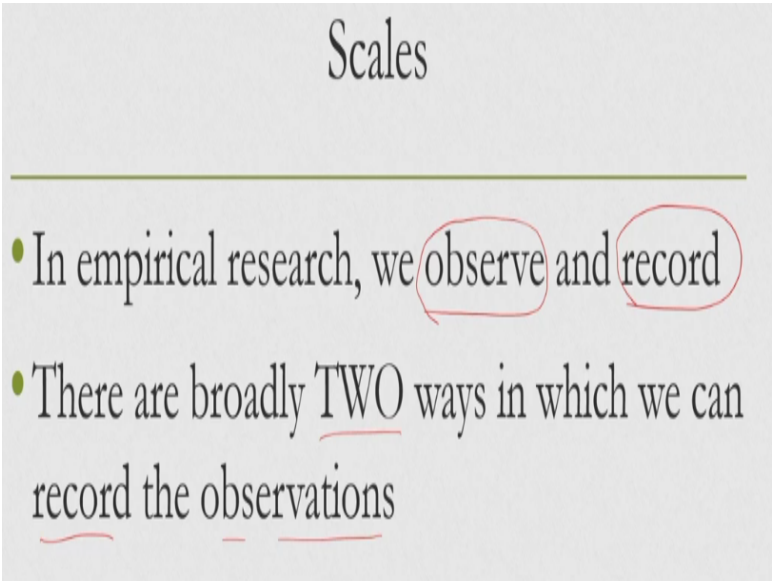
(Refer Slide Time: 10:30)



And once we are able to identify different test conditions in a systematic manner, we can perform the empirical study. So, empirical studies are done based on these test conditions. In other words, identification of variables is necessary to be able to perform the study. Along with the variables, it is also important to suitably define their values. So, what kind of values they can take that also needs to be carefully decided and defined.

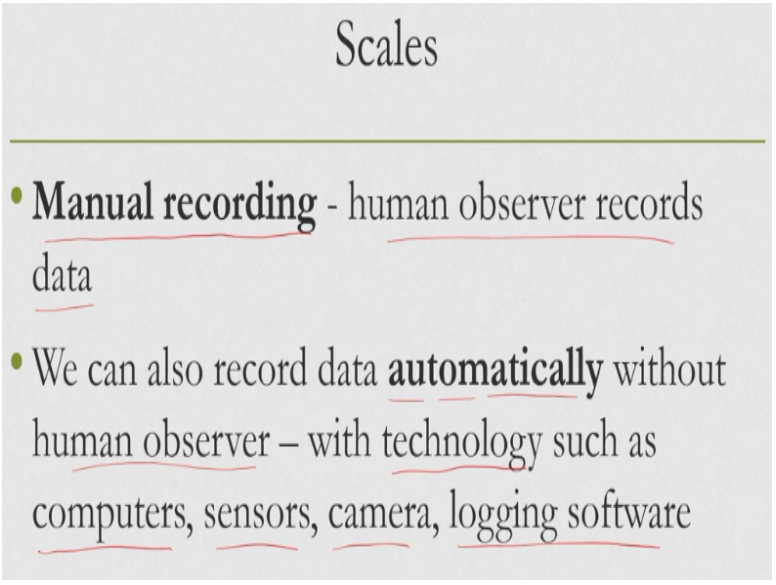
This brings us to the idea of measurement scales. We need to have some knowledge of measurement scales to be able to suitability define the values of the variables that we are going to use in our study.

(Refer Slide Time: 11:22)



So, what is this idea of measurement scales? In empirical research as we have repeatedly mentioned, overall objective is to observe, in our case to observe user behaviour and record the behaviour by some means. Now, there are broadly two ways in which we can record the observations, what are those two ways?

(Refer Slide Time: 11:48)

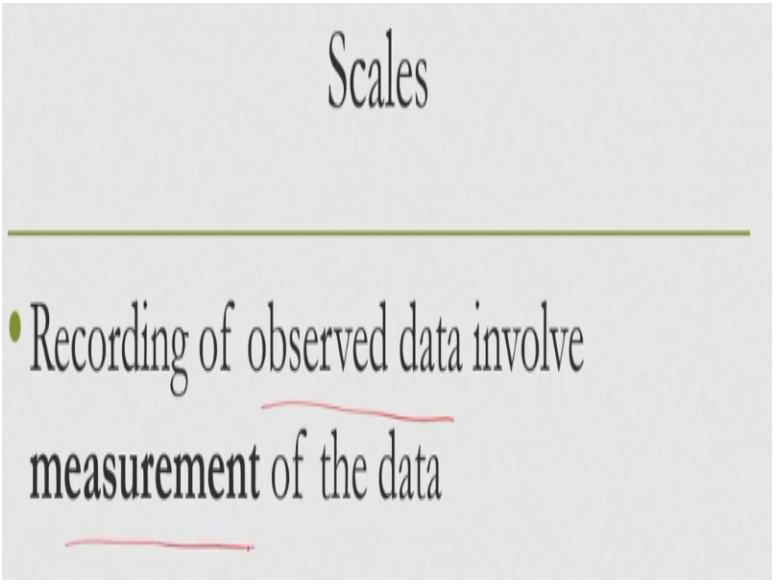


One can be manual recording that means there is a human observer who is actually there in the experimental setup where the study is taking place and manually recording the observed data. Other can be automatic recording that is we can also record the observed data

automatically without the intervention of any human observer, how we can do that? With the help of technology such as computers, different sensors, camera, logging software, etc.

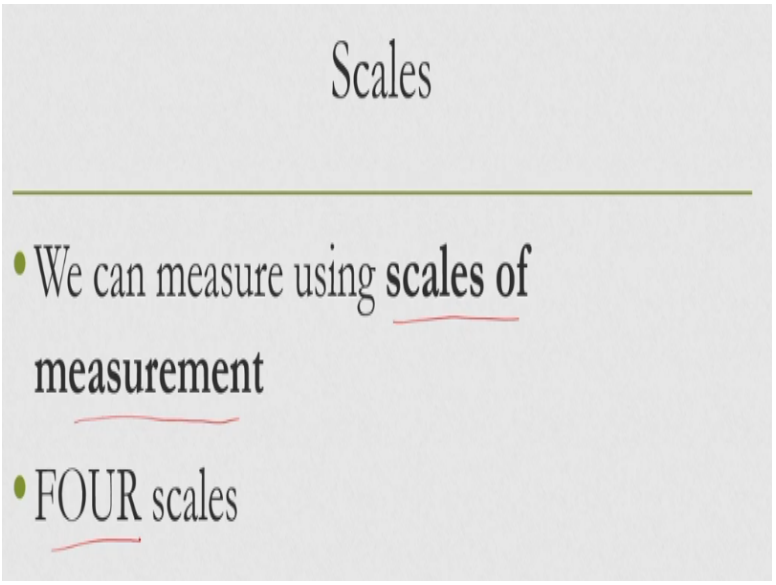
So, these can also be used to automatically log or record the behavioural data instead of somebody sitting there and manually recording the data. So, these are broadly two ways, manual recording and automatic recording.

(Refer Slide Time: 12:48)



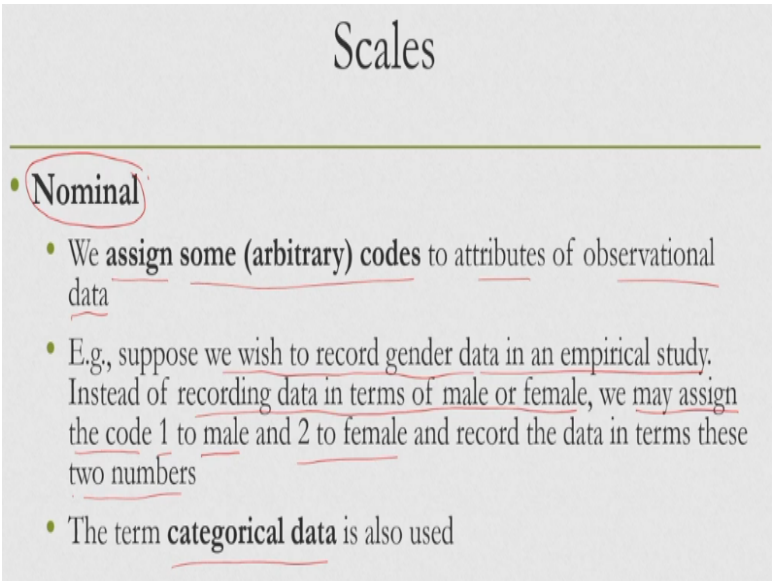
Now, when we talk of recording of observed data, it involves the idea of measurement of the data. How do we measure and represent it?

(Refer Slide Time: 13:02)



We can do so, we can measure data using scales of measurement and there are broadly four such scales that are available.

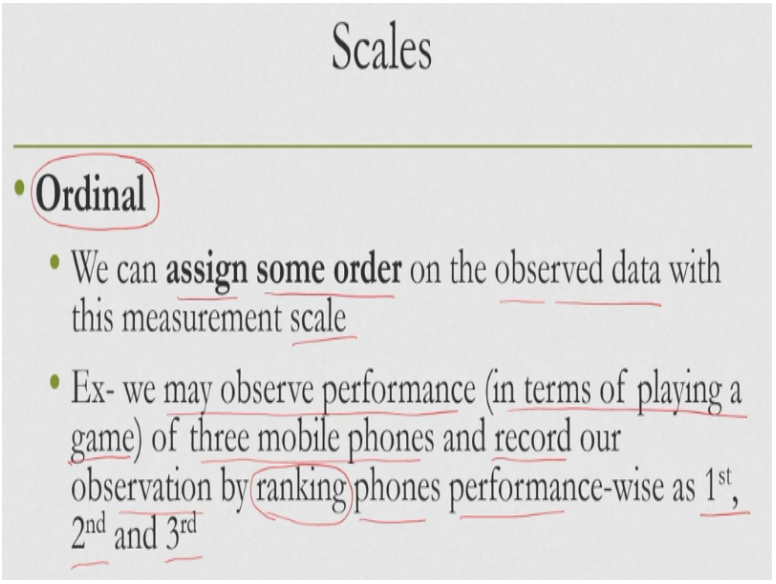
(Refer Slide Time: 13:14)



One is the nominal scale of measurement, what this scale implies? If we are going to use a nominal measurement scale, then what we do is we assign some arbitrary codes to attributes of observational data. So, some arbitrary code is assigned to the data for the purpose of recording the data. For example, if we wish to record gender data in an empirical study, instead of recording data in terms of male or female, what we can do is we may assign the code 1 two male and 2 female.

Again, this is a very arbitrary assignment, so there is no particular logic inherent, but we can do so, we can assign the code and then record the data in terms of the code, these two numbers. So, whenever male needs to be recorded; we record it as 1, whenever female needs to be recorded we record it as 2. The nominal scale of measurement also occasionally termed as categorical data. So, whenever we are using nominal measurement scale and recording data that data is sometimes called as categorical data.

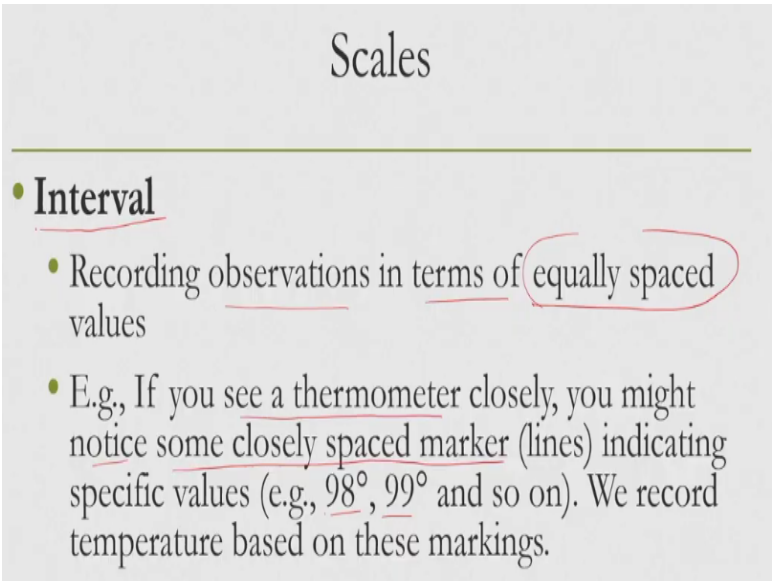
(Refer Slide Time: 14:41)



Next scale of measurement is ordinal scale. Here what we do? We can assign some order on the observed data with this particular scale. Let us try to understand this within an example. We may observe performance say in terms of playing a game, it has defined the idea of performance as a game playing experience.

So, we may observe performance of 3 mobile phones and record our observation by ranking, introducing a rank among the 3 three phones by ranking those phones, that is their performance in the form of first, second, third, etc. So, this first, second, third this type of recording is using ordinal measurement scale.

(Refer Slide Time: 15:47)

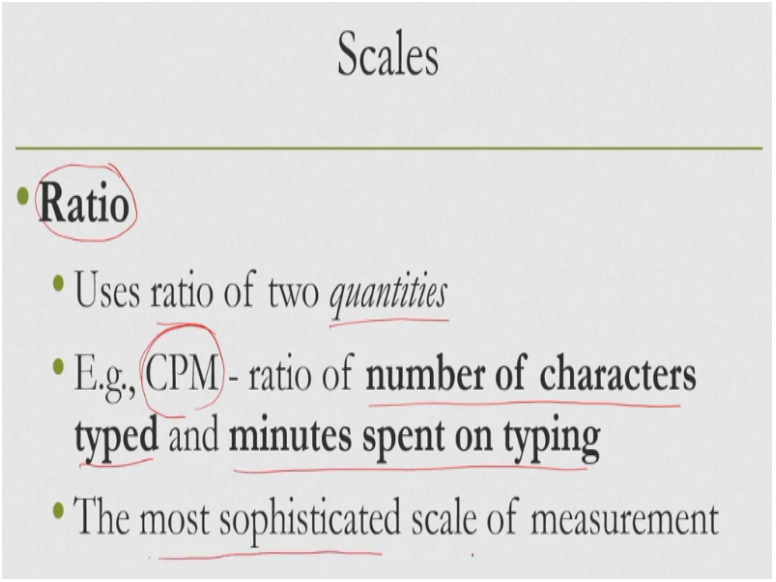


Next is the interval scale of measurement. So, here what we do is we record observations in terms of equally spaced values. So, in the interval scale of measurement, we use equally

spaced values to record our observations. For example, if you see a thermometer closely, you may notice that there are some closely spaced marker or lines, each of these lines indicate some value, for example 98 degree, 99 degree and so on.

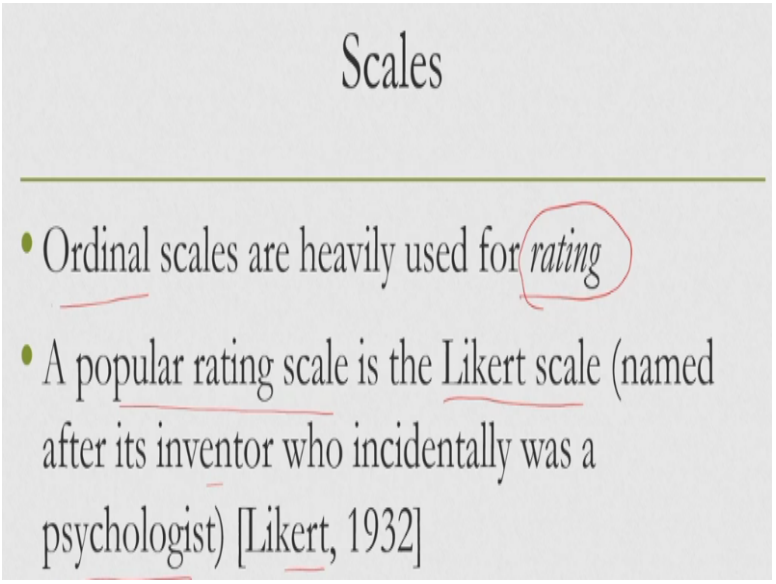
So, the gaps between those values are same, so they are equally spaced. And if we are using this to record our observation that is the temperature then we are actually using the interval scale of measurement that is the third scale of measurement.

(Refer Slide Time: 16:36)



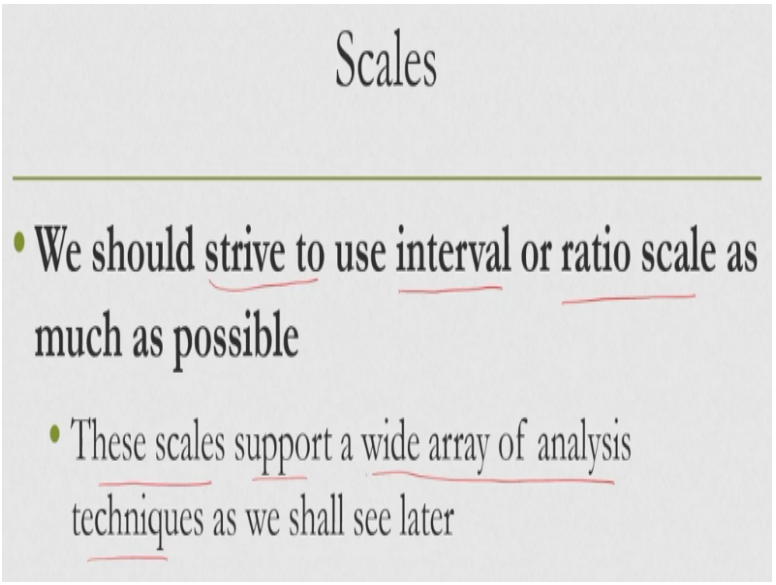
We also have a final and most sophisticated measurement scale that is ratio scale of measurement. So, this scale uses ratio of two quantities to record observations. For example, the idea of CPM or character per minute, so this is a ratio of two quantities, number of characters typed and the minutes spent on typing. So, if we take the ratio of these two quantities, then we get the data and here we can say that we are using ratio scale of measurement. Now, ratio scale is considered to be the most sophisticated scale of measurement.

(Refer Slide Time: 17:32)



It will be useful to note that the ordinal scale is heavily used for rating. So, occasionally we need to rate various aspects of a design and the ordinal scale is primarily used for that purpose. A popular rating scale is the Likert scale, it is named after its inventor Likert which was invented way back in 1932 and the inventor was a psychologist who came up with the idea of this scale. Likert scale is a very popular scale that is used for rating. So, there are four scales of measurement; nominal, ordinal, interval and ratio.

(Refer Slide Time: 18:05)

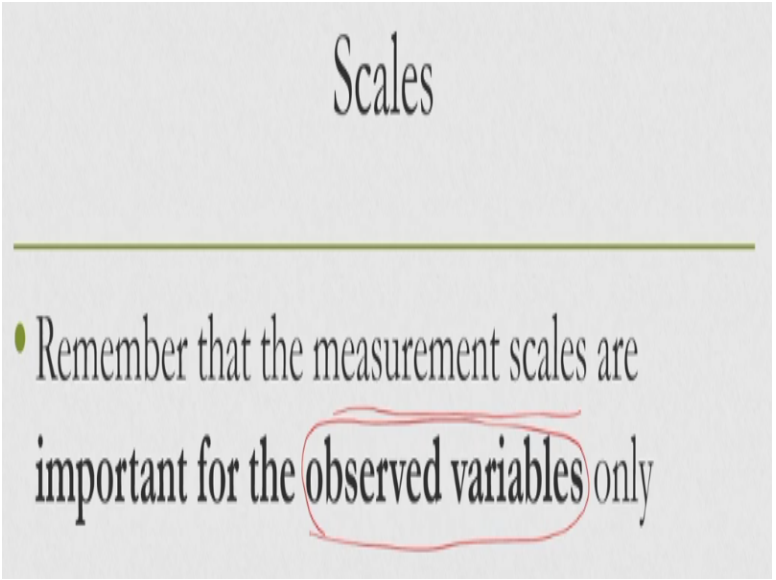


Now, what should be our objective or what should we strive for while trying to choose a particular scale of measurement? We should strive to use interval or ratio scale as much as possible. So, whenever we are trying to design some experiment and need to record data, our objective should be to design the experiment in such a way that we can record data either using interval scale or ratio scale that will give us a better way of recording data.

Why we need to go for these types of scales? These skills support a wide array of analysis techniques as opposed to the other scales. So, one objective of collection of data is to be able to analyse them so that we can come to a reliable conclusion about the observations about the system. If we use interval or ratio scale, then we will be able to employ a wide array of analysis techniques on the data which can help us get to a reliable conclusion.

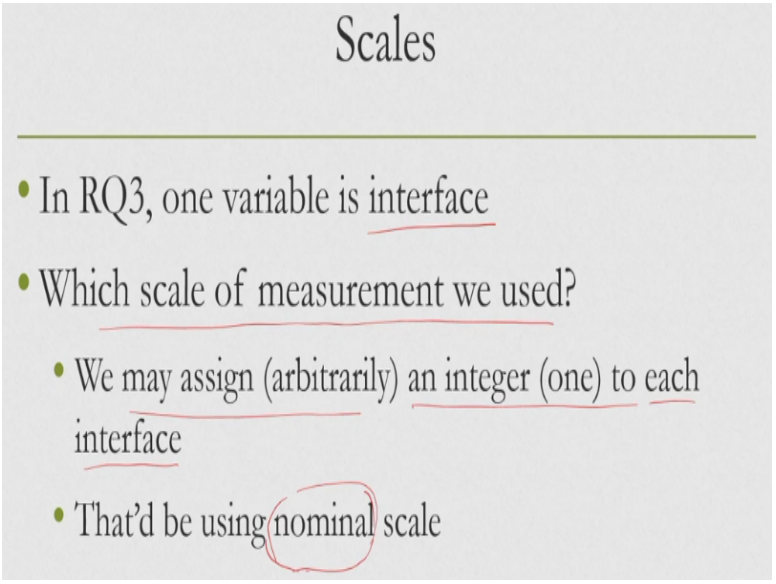
If we use the other scales, then we may not be able to use those techniques, so it will be difficult to conclude. So, we should strive for using either of these two scales, ratio scale is even better if we can manage with only ratio scale.

(Refer Slide Time: 19:45)



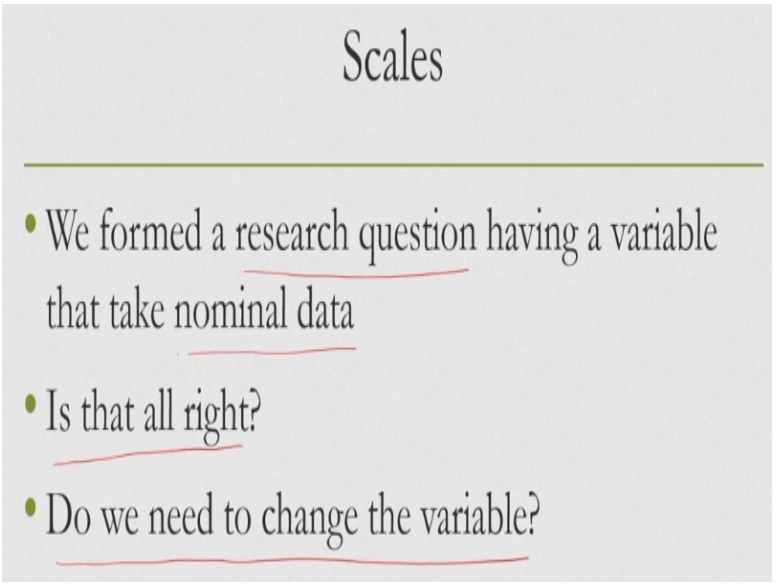
Now, one more thing that you should note here is that the ideas of the measurement scales are important for the observed variables only. In fact, when we; are talking about variables, different types of variables there. Now, the measurement scales are not required to be considered for variables that are not observed that are controlled, we need to consider measurement scales for observed variables only.

(Refer Slide Time: 20:13)



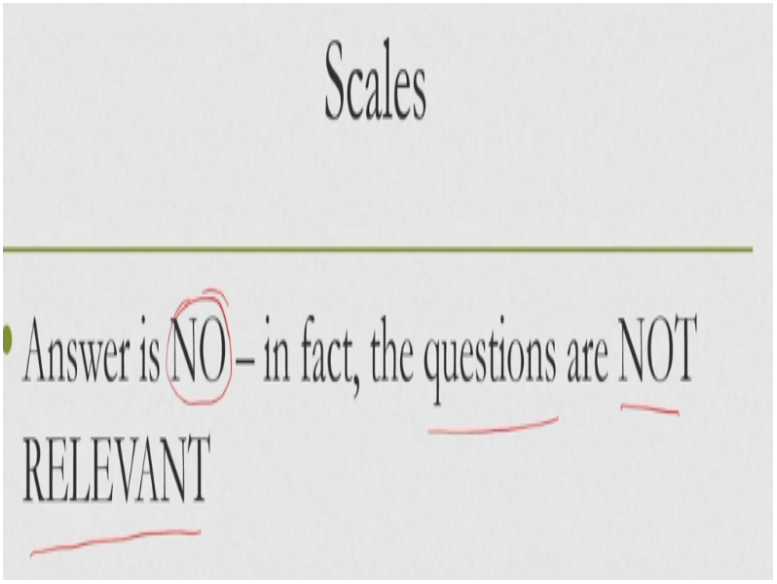
For example, let us reconsider the research question 3 that we have seen earlier. Now, one variable there is the interface variable. It may come to you that for this variable; which scale of measurement we should use? What should be the most appropriate scale of measurement? Remember that interface variable takes value as different instances of typing interfaces. Now, we may assign arbitrarily an integer to each interface. In other words, we will be using the nominal scale of measurement as per definition of the nominal scale.

(Refer Slide Time: 20:55)



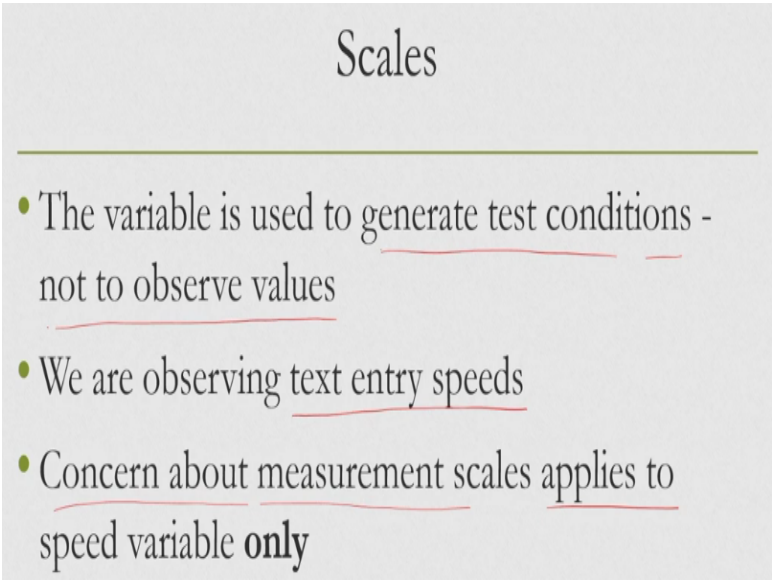
Then, we formed a research question having a variable that take nominal data. So, that is the situation we are in if we choose to use nominal scale of measurement for interval variable. Now, is that the right thing we did? Do we need to change the variable? These are maybe some questions that may confuse our mind that may agitate our mind.

(Refer Slide Time: 21:24)



But the answer to both the questions is a big no. In fact, the questions themselves are not relevant at all, we do not need to bother about the measurement scale for the interface variable, why?

(Refer Slide Time: 21:42)



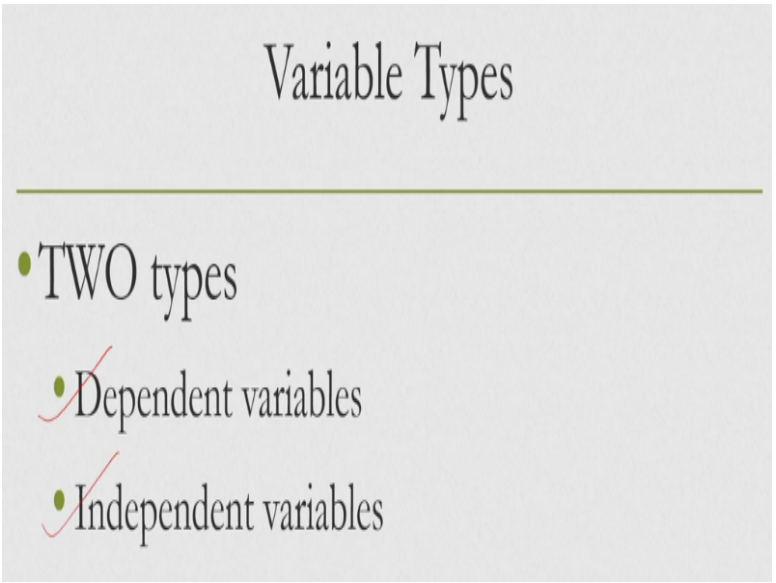
The interface variable is used to generate test condition that is we are assigning it different values so that we can set up different experimental conditions. If we assign it the value of our interface, our design; then we have one test condition where users will type with our design. If we assign the interface the value of MS Word, then the users will get to type with MS Word. So, we are essentially using the variable to generate test conditions by assigning different values, but these assignments have nothing to do with observation of the values.

So, we are not going to observe anything, we are simply assigning So, interface is not an observed variable. What we are observing in this case? We are observing text entry speeds. So, text entry speed is the observed variable, not the interface. So, whether we use any measurement scale or not for interface is irrelevant, we need not bother about that. What we should bother about is what kind of measurement scale we should use for the observed variable here that is the text entry speed.

So, the concern about the measurement scales applies only to the text entry speed variable. Those are the basic ideas about variables. What is a variable, why we need it, what kind of variables are there. How to decide on measurement scales for a variable. Which scale is good and for which type of variable the scales apply. Now, let us learn a little bit more about different types of variables.

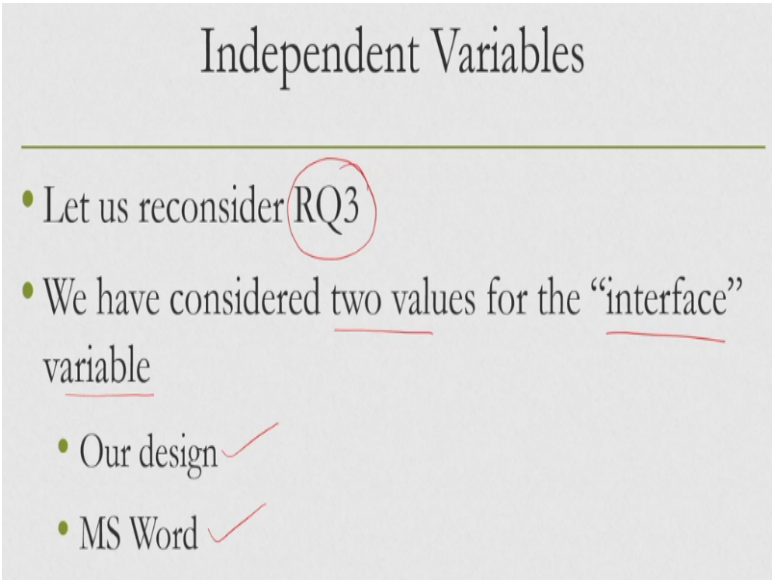
Earlier we have broadly categorized them as observed variable and non-observed variable, but that is not the right term to use, so we will use more popular frequently used terms. The first such terms that we should know about are dependent variables and independent variables.

(Refer Slide Time: 23:49)



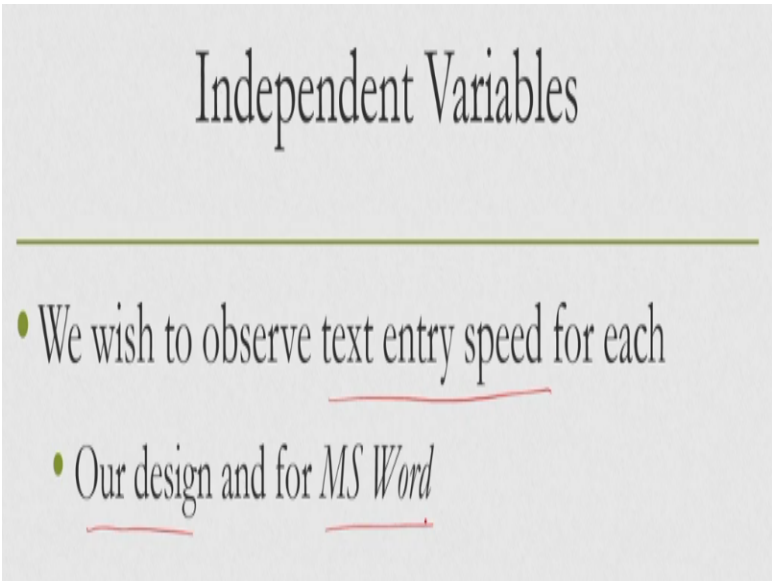
So, variables are broadly of two types, one is called dependent variable, other one is called independent variable. What are these variables let us try to understand.

(Refer Slide Time: 24:01)



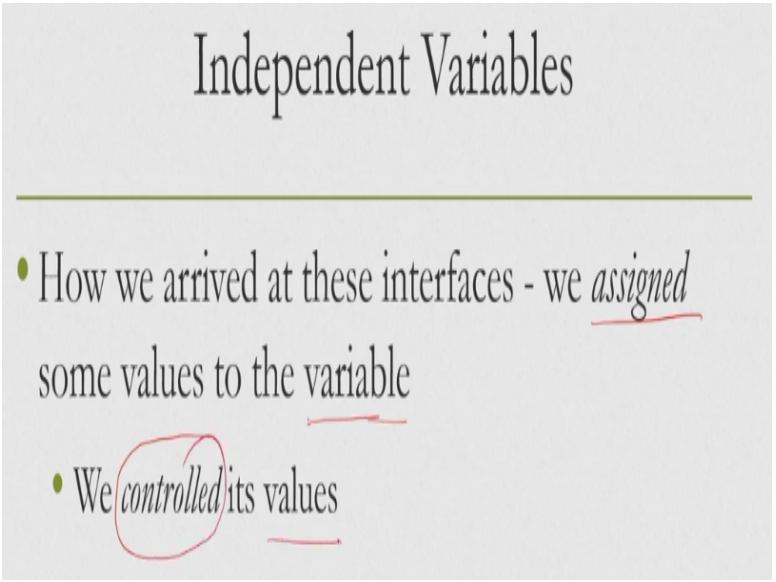
Let us reconsider the research question 3. Now, we have considered two values for the interface variable as we have mentioned earlier. One value is the; our design that is the GUI that we have designed for text input, other value is MS Word an existing text input interface. So, these two values we have used to instantiate the interface variable.

(Refer Slide Time: 24:32)



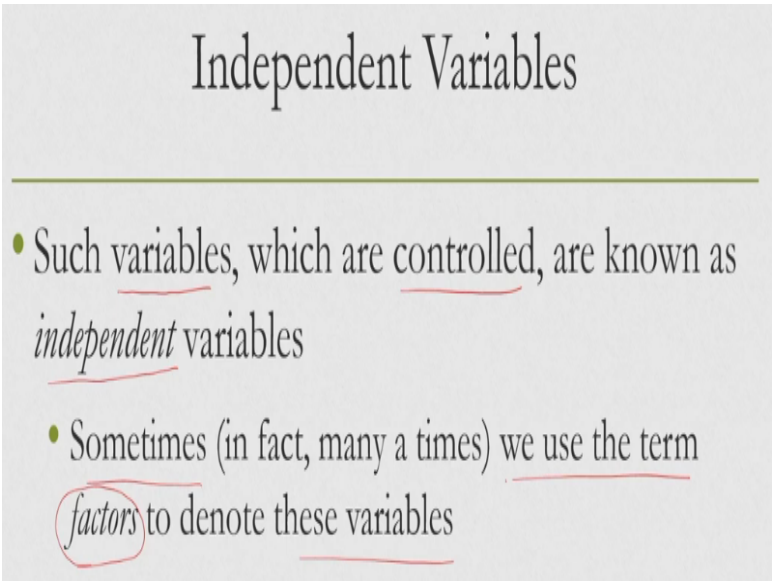
What is the purpose? We wish to observe text entry speed for each of these interfaces that are for our design and for the MS Word interface.

(Refer Slide Time: 24:45)



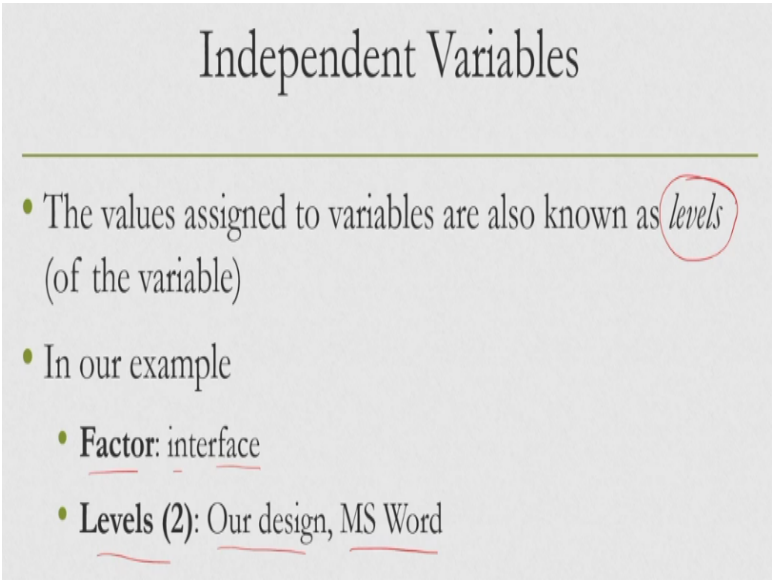
Now, how we have arrived at these interfaces? How we have chosen these two values? We assigned some values to the variable. By assigning values to the variables, we have arrived at these two interfaces on which we are going to perform the study and observe the text entry speed. So, essentially what we did? We controlled in a manner its values.

(Refer Slide Time: 25:12)



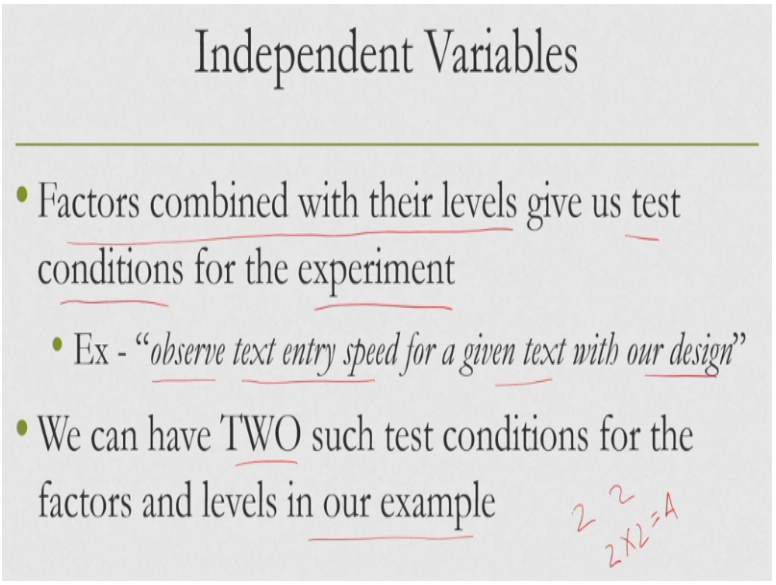
Such variables which are controlled by us, the experiment designers are known as independent variables. So, the interface is an independent variable. Sometimes, in fact quite frequently we use the term factors to denote these variables. So, either we can use the term independent variable or we can use the term factors to denote variables that we are controlling in order to generate different test cases.

(Refer Slide Time: 25:52)



So, when we are assigning values to independent variables or factors, those values are also known as levels. So, if we use this terminology then in our example of the research question 3, RQ 3 we have one factor that is interface and we have two levels that is our design and MS Word. So, each value is a level and the factor name is interface.

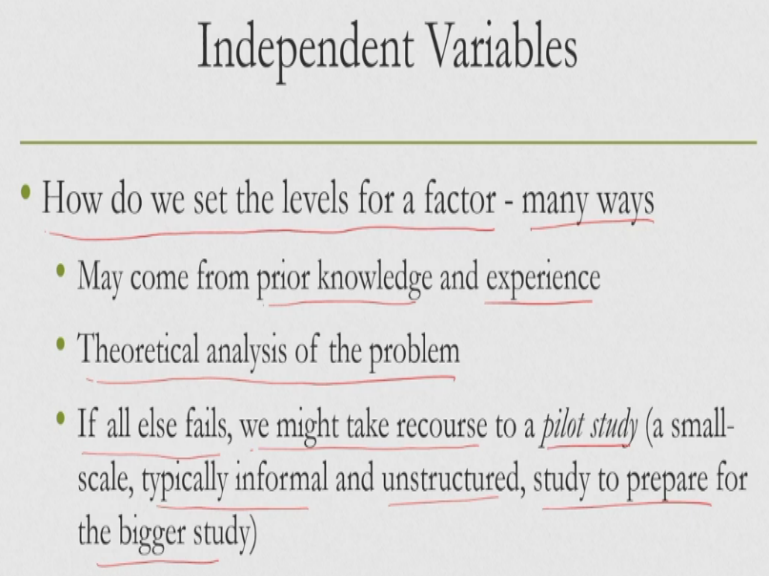
(Refer Slide Time: 26:24)



When we combine the factors with levels, we get the test conditions. So, factors combined with levels give us test conditions for the experiment, which is the primary objective of identification of variables. For example, observe text entry speed for a given text with our design. So, here we are combining the factor that is interface with the level that is our design to generate this test condition that is observe the speed for a given text with our design.

Since, we have one factor into levels, we can have two such test conditions in our example. So, if we have supposed two factors and each factor can have two levels, then how many test conditions we can generate, 2 into 2 equal to 4 such test conditions we can generate. So, by multiplication we can get the number of test conditions that we can generate, multiplication of factor and levels.

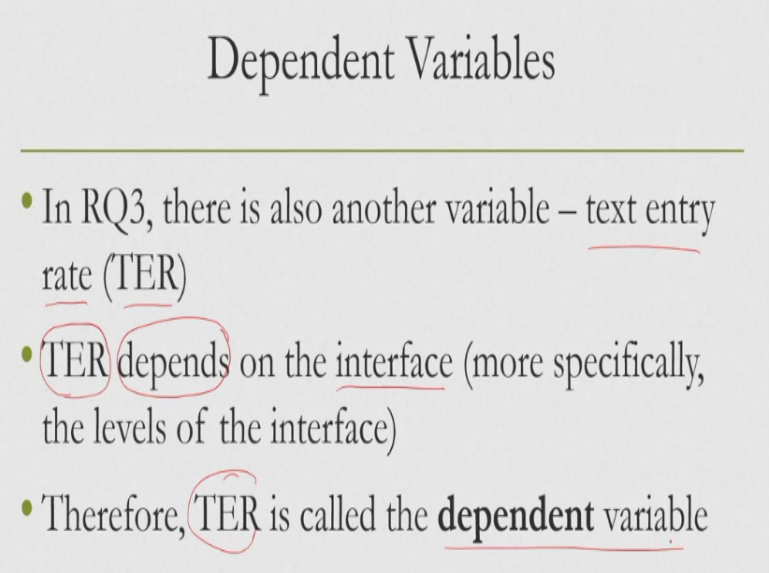
(Refer Slide Time: 27:36)



Now, the next question that may come to our mind is how do we set the levels for a factor? How do you decide what are the values are to be assigned to an independent variable? There can be many ways to do that. Such assignment may come from prior knowledge and experience, we may already have some experience of doing that. Theoretical analysis of the problem can also reveal that.

If all else fails, we might take recourse to a pilot study which is a small scale, typically informal and unstructured, study to prepare for the bigger study. So, we may conduct a pilot study to identify the levels of a factor that are going to be suitable for the particular study that is about the idea of independent variables and its values.

(Refer Slide Time: 28:29)

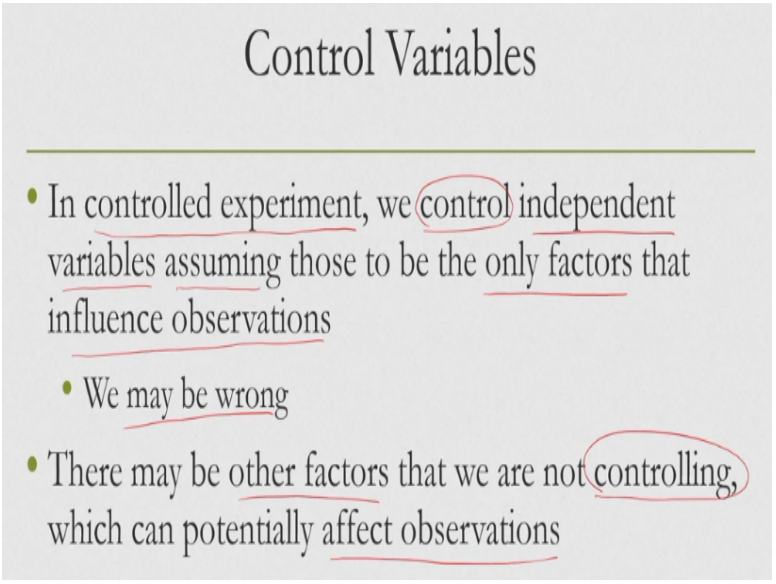


The other type of variable that we mentioned is dependent variable. Now, in the research question that is RQ3 another variable was used that is text entry rate or TER to decide whether one is faster than the other. Now, this text entry rate variable depends on the interface. In other words, as we change the levels of the interfaces, the rate is expected to change. Since it depends on the independent variable that is the factor, we call TER as the dependent variable, dependent on the independent variable.

So, that is the idea of dependent and independent variable. So, when we control the values to generate test conditions, we control the values of a variable we call that variable as independent variable. Often, such variables are called factors and the values they take are often called levels. Depending on the number of factors and number of levels in each factor we can compute by multiplying the total number of test cases that we can test, scenarios that we can generate for our study.

And there is another category of variable that is the dependent variable which depends on the factor and levels that is one set of terminology. Another set of terminology we should be aware of is the idea of control variable and confounding variable. Let us see what are those variables.

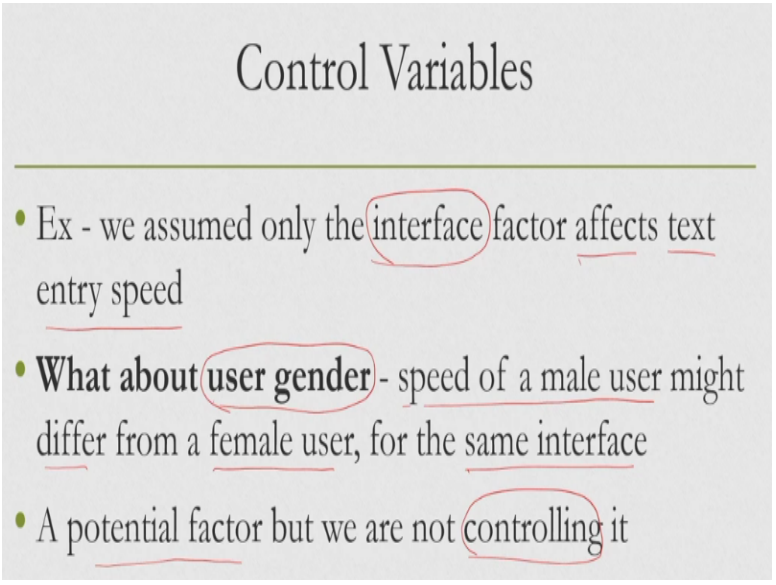
(Refer Slide Time: 30:17)



Let us start with the idea of the control variable. In a controlled experiment, we already mentioned that we control independent variables, this is very important, assuming those to be the only factors that influence observation. So, when we are setting up an experiment, we generally assume that the values of the independent variables that we are controlling are the only thing that needs to be controlled.

There are nothing else and these control values only influence the outcome, but we need not be entirely correct, we may be wrong. There may be other factors that we are not controlling, which can potentially affect the observations, which can influence the behaviour and affect the observations.

(Refer Slide Time: 31:15)

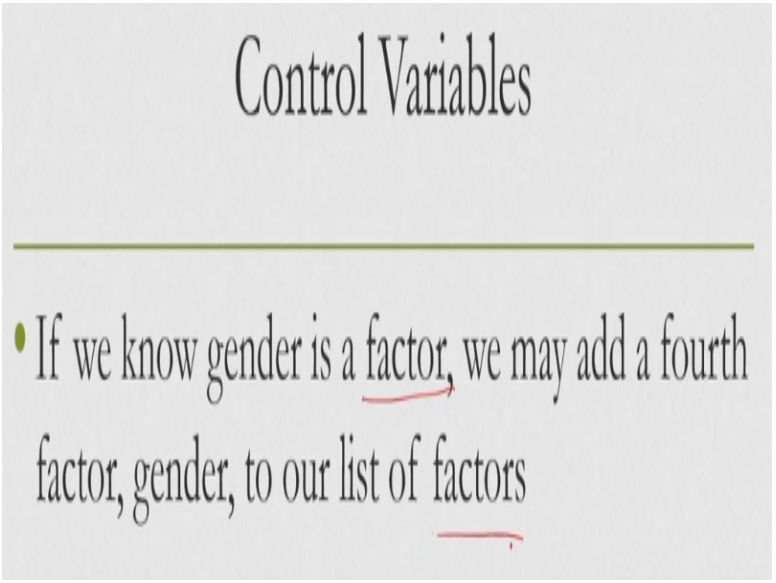


For example, in our earlier research question 3, we assumed that only the interface independent variable or factor affects the dependent variable that is the text entry rate. But then what about another factor say user gender, speed of a male user may differ from a female user, quite possible for the same interface, we have no knowledge and we have not explicitly taken into account this factor.

So, it can be a potential factor, but we are not controlling it. So, later on the data that we record may be biased, how? Say in our experiment we only considered male participants to collect data vis-a-vis conducted another experiment where we considered majority of male participants and few female participants. In a third way, we can consider majority of female participants and few male participants and in another case we can consider only female participants.

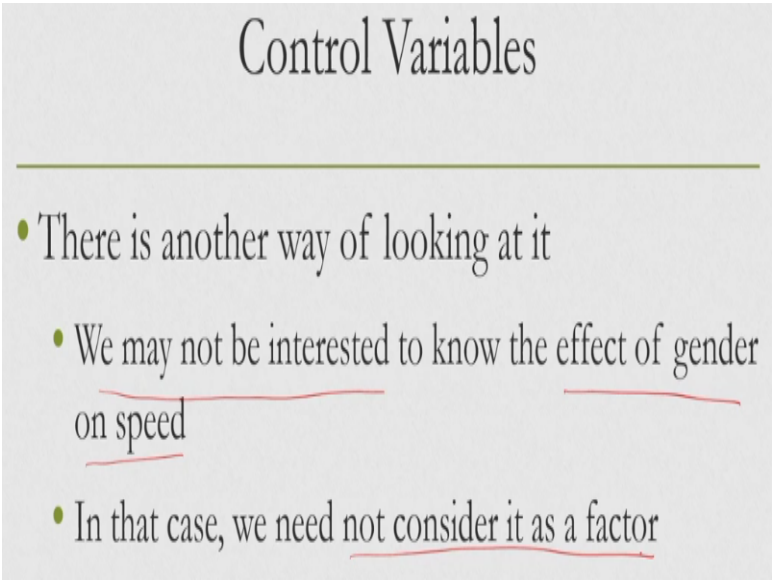
We do not know what kind of participants we are going to use because we have no idea that the gender of participants can affect the potential outcome. So, we may miss this factor and its influence on the outcome.

(Refer Slide Time: 32:41)



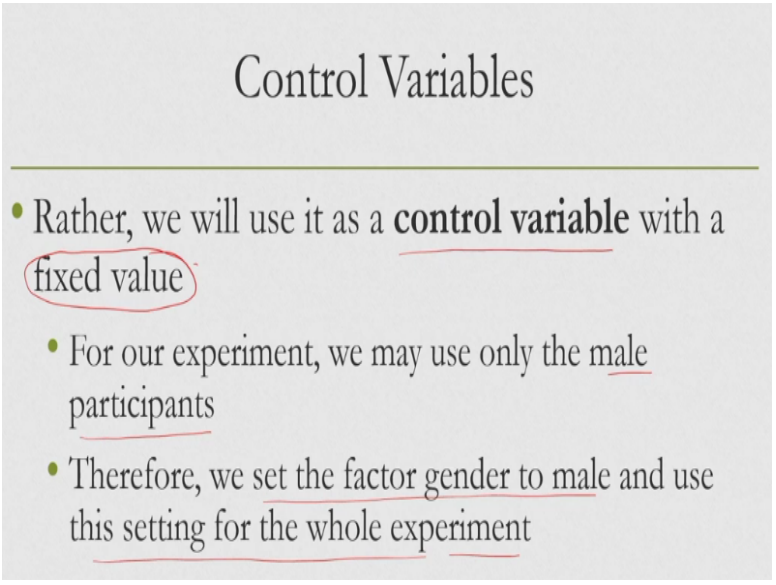
But if we know gender is a factor, then we may treat it as another independent variable and add it to our list of factors that is possible.

(Refer Slide Time: 32:52)



There is another way of looking at it. We may not be interested to know the effect of gender on speed, although we know that gender may affect the speed, but we are not interested to know how it affects unlike the earlier case where we are interested to know how a design affects the speed. So, in that case what we can do? We need not consider it as a factor as a variable, we need not vary it.

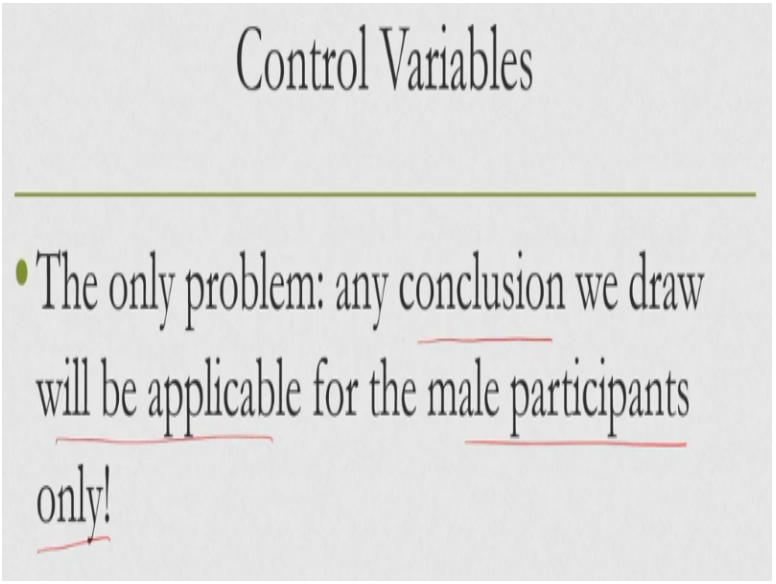
(Refer Slide Time: 33:24)



Rather, what we can do is we can use it as a control variable with a fixed value. Although we are using the term variable, we do not vary it, we keep it fixed. For example, in our experiment what we can do is we may use only the male participants or only the female participants, but not a mix of them. So, in that case we are fixing the value of gender to a fixed value that is either male or female.

Therefore, we set the factor gender to male if we use only male and use the setting for the whole experiment, so essentially we are treating this variable as a constant, although we are calling it control variable, but we are treating it as a constant value.

(Refer Slide Time: 34:11)

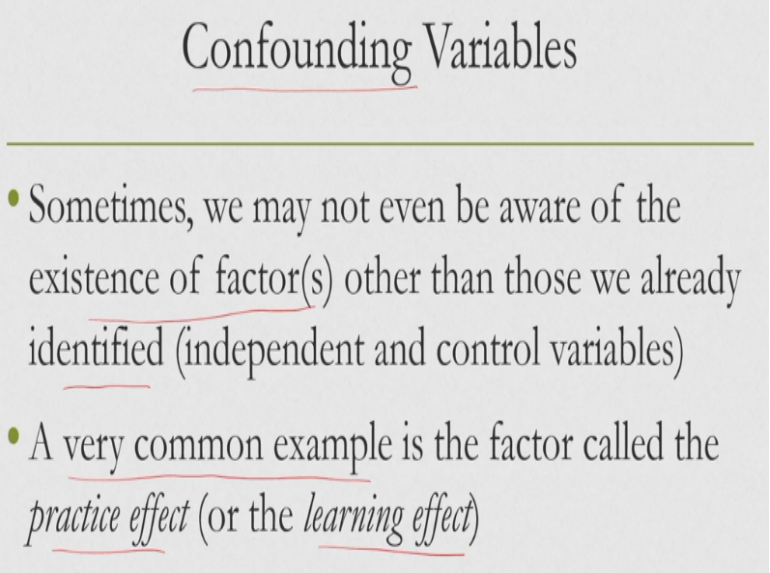


Now, this can solve our problem of not complicating the experiment, but this introduces another problem which we should be aware of that is any conclusions that we draw based on the data that we collect from only male participants will be applicable for male participants only. So at the end of the experiment, we collected some data about this speed and then decided to analyse it.

And concluded that our design is faster than MS Word based on data collected from only male participants then our conclusion will be applicable only for male participants. It may change for female participants because in our experiment we are not considering gender as a variable rather we are fixing the value of the gender value. So, conclusion will also be accordingly applicable only to the fixed value of the control variable that is about the idea of control variable.

Where although it may be an independent variable or a factor, we are not treating it as a variable explicitly, instead we are fixing a value of this variable and using that value throughout our experiment. So, in other words we know that that can be affected, but we are not interested to know the relationship between that factor and the outcome.

(Refer Slide Time: 35:13)

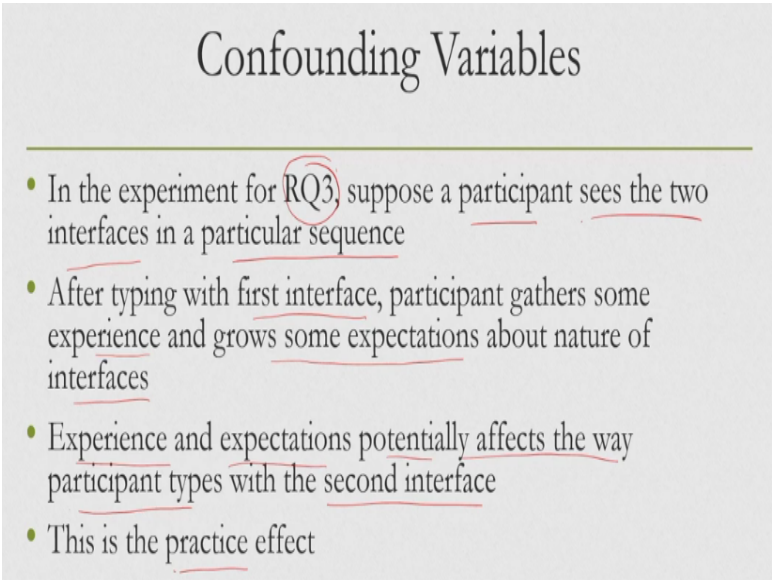


Another group of variables is known as a confounding variable. Sometimes what may happen is that we may not even be aware of the existence of factors other than those we already identified. So, earlier we said that there are factors that affect the outcome that is the dependent variable. First, we learned about independent variables that affect the outcome. Later on, we learned about control variables that may affect outcome.

So we are fixing their values, all these assume that we are already aware of them. Now, it may so happen that we are not even aware of presence of some variables that may affect outcome, so then what will happen in that case? The outcome that we get without knowing that there are other variables at play may not be the exact outcome that we are looking for. Let us see one example.

A very common example is the factor called the practice effect or the learning effect. This is one such variable whose presence we often do not recognize, but it affects the outcome. What is this?

(Refer Slide Time: 36:50)

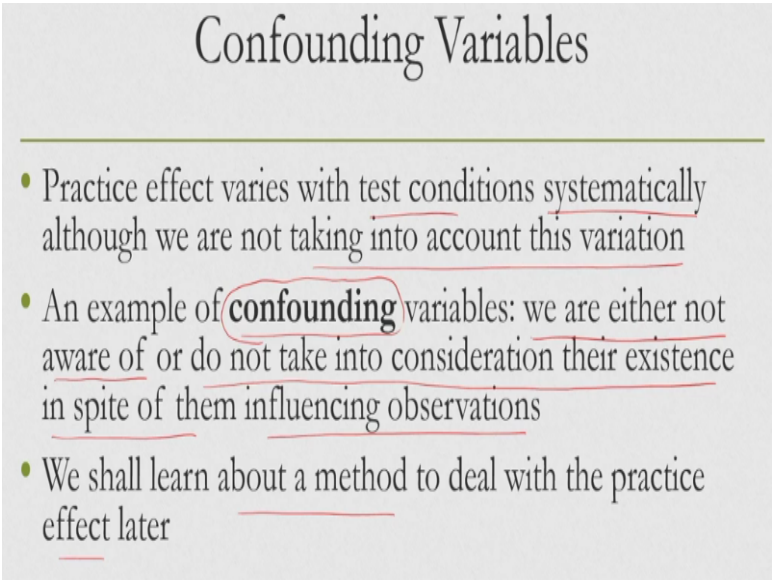


In the experiment for RQ3, research question 3, suppose a participant sees the two interfaces in a particular sequence. After typing with the first interface participant gathers some experience of typing with the first interface and grows some expectations about the nature of the interfaces because some tasks have been performed with the first interface. Now, this experience and expectations potentially affects the way the participant types with the second interface.

Now, while typing with the second interface, the participant expects that the feature should be at a particular place or the layout should look something like this or the icon should be something like this because already the participant has experienced similar things with another interface. So, that expectation may affect the performance, participant may spend some time looking for such things, then realizing that this is a different thing, so eventually his or her speed of entry may get affected. So, this is called the practice effect.

So, we practiced with one thing and gathered some experience with that thing, then next when we are asked to do or perform some tasks with some other interface, then the effect of the practice with the first interface may affect the outcome with the second interface, this is the practice effect or the learning effect. We have learned the first interface and expect the second interface to be similar to the first one, although they are different and this affects the outcome.

(Refer Slide Time: 38:43)



Now, practice effect varies with test conditions in a systematic manner, although we are not taking into account this variation. So, when we design the experiments, we often miss to recognize that there is this thing called practice effect and we design test conditions in a way where the practice effect may vary systematically. So, this is an example of a different type of variable called confounding variable. What are these variables?

We are either not aware of these variables or do not take into consideration their existence in spite of them influencing observations. So, in case of control variables we know that they exist and we control them by fixing their value. In case of confounding variables, we are not aware of their existence or are not taking explicitly into account their presence in spite of those variables affecting the outcome.

Of course, the question may come to your mind that if such is the case, if they are confounding variables and we are not taking care of those variables, then how reliable will be the data, the observation because as I said confounding variables may not be known to us, so how do we take care of it? There are some common confounding variables such as the practice effect, we already know they exist.

So there are ways to take care of those variables and in subsequent lectures we shall learn about a method to deal with the practice effect. But, of course, in general if we are not aware of the presence of any arbitrary confounding variable, then it will be difficult to take care of it and the conclusions that we draw may be biased and may not lead to a reliable conclusion. So, in this lecture we have learned few things.

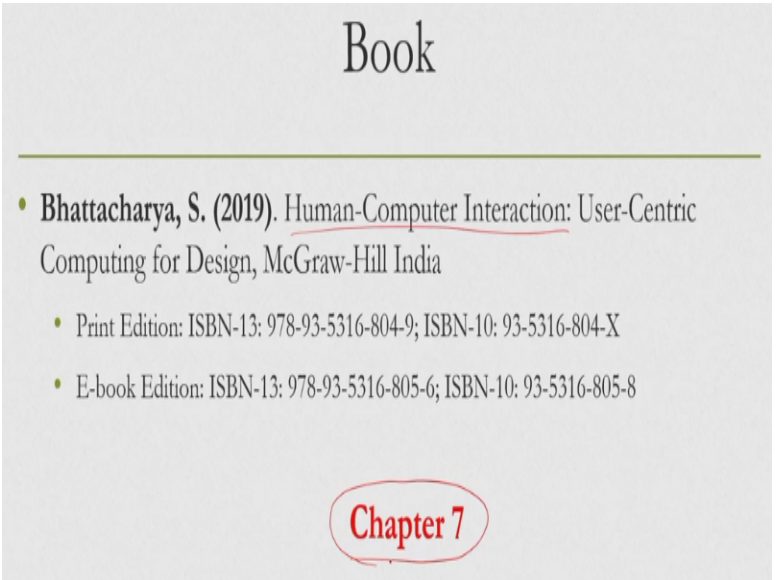
We are discussing the second stage of the empirical study that is identification of variables. So, what we have learned? First of all what is a variable, why it is needed? Then we learned about measurement scales and why that is important. So, there we broadly saw that we need to bother about measurement scales only four variables whose values we are observing rather than variables whose values we are not observing.

Now, broadly we have divided variables into two groups, dependent variables and independent variables. Independent variables are also known as factors and the values that they take are known as levels. Whereas dependent variables are those that are dependent or the value of those variables are dependent on the independent variables and their values. We have also talked about control variables.

That is there are some variables which we are aware of which we know that they are present and they affect the outcome, but we are not interested to know about how they affect outcome. So, in that case, we simply fixed their values, we treat them as constants, those are called control variables. And also, we discussed about an important class of variables, confounding variables.

These variables are more difficult to handle, we may not be even aware of their presence. However, there is a very common confounding variable that is called practice effect or learning effect. And later on, we will see how to deal with such a variable.

(Refer Slide Time: 42:39)



Whatever I discussed in this lecture can be found in this book Human-Computer Interaction, you can refer to chapter 7. I hope you have enjoyed the content and understood the concepts that we covered today in this lecture. We are going to continue our discussion on the empirical study, the other stages of the study, namely the experiment design stage and the data analysis stage. That is all for this lecture. Thank you and goodbye, looking forward to meet you all in the next lecture.