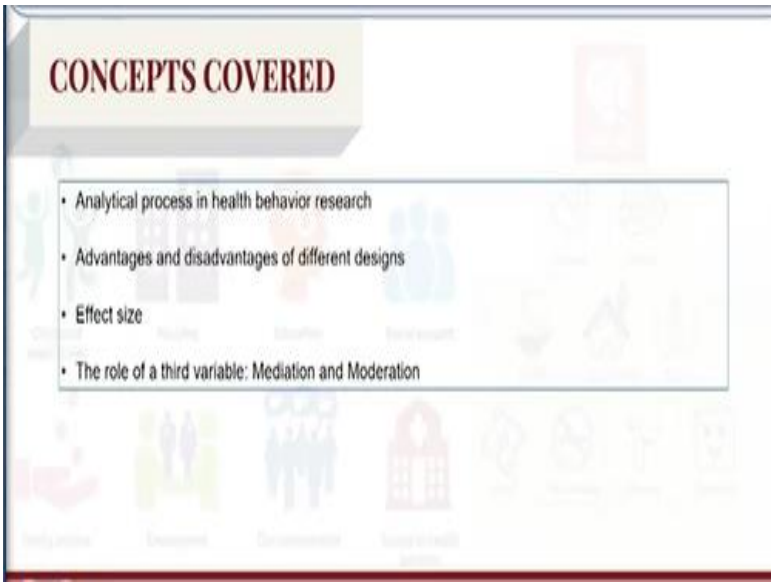
Basics of Health Promotion and Education Intervention
Dr. Arista Lahiri
Dr. B C Roy Multi Speciality Medical Research Centre
Indian Institute of Technology, Kharagpur

Lecture - 38
Analyzing Health Behavior Change Data

We are in the pursuit of understanding the basics of health promotion and education intervention. And in that pursuit, we have ended up in the last week of it and we were now discussing about how to evaluate the health behaviour change interventions. Because planning is done the process is done that is the implementation is done. We have the concept we have the idea we have done it. Now we want to evaluate whether our intervention is effective or not.

So, in this lecture after the discussion on evaluation models and the types of evaluation we now move on to the analyses techniques like we will be discussing about how to analyse all those health behaviours change data that we have gathered through the implementation stage of our program.

(Refer Slide Time: 01:04)



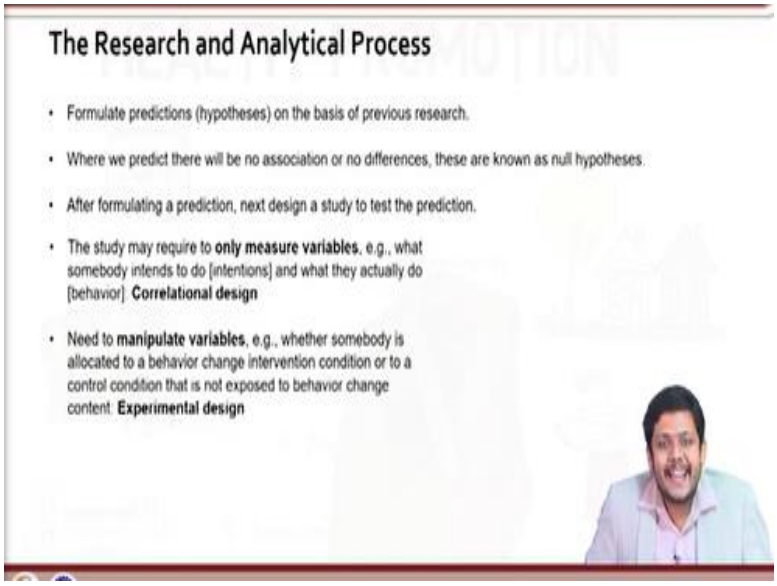
So, for that we have the analytical process in health behaviour research. We will be covering the advantages and disadvantages of different designs, effect size this is very important because this is an epidemiological concept or a statistical concept. But in fact, I would say it is a holistic

analytical concept. Whenever you do any kind of facilitated data analysis or in any other data analysis you come across effect size.

This is how you express how the factor is causing the outcome like that and then the role of a third variable. We were discussing about the contextuality we were discussing about mediation and moderation now finally. Now we will be coming across the mediation and moderation part and will be discussing about mediators and moderators.

(Refer Slide Time: 01:50)

(Refer Slide Time: 01:51)



So, let us directly dive in now this is interesting. Regarding the research and the analytical process for this health promotion and education intervention part what we must understand is we must first understand the concept of hypothesis that is here we formulate a prediction and then we test it. So, the prediction that we formulated is the hypothesis and whatever we are testing we are testing the hypothesis.

When we are testing when we are I mean discussing the hypothesis in a form that, no this behaviour is not bringing any change that is called a null hypothesis or hypothesis of no difference. And in the other part of it or against it stands the alternative hypothesis through which we state that, no it did have some changes. Now when we say like this that it did have some changes, I mean what happens is it brings about two directions, change for good and change for bad.

This is in a statistical way called two tails. So, alternative hypothesis again can be either two-tailed or one-tailed and one-tailed means in one direction. So, then it can be like if I am considering the null hypothesis that I have given a behaviour change intervention through a mobile app for quitting smoking. What can be the outcome the person has quit smoking? The outcome can be that the person did not quit smoking.

The person under that did not quit smoking can be I mean of two types that smoking the quantity decreased that is again a good factor and the quantity increased. So, the null hypothesis is the person did not quit smoking or there is no difference from the previous part pre-intervention phase and the post intervention phase. But the alternative hypotheses are two types like one is the behaviour continued and the behaviour stopped.

Here the behaviour stopped is our more desired outcome. So, now when we are testing this hypothesis when we are rejecting the null hypothesis, we have to identify the direction to which our test results are going. So, through that we can then accept the alternative hypothesis. In a simpler way we can say that if there is any difference that there is any change, we can say that the null hypothesis is excluded.

But again, these are just a brief discussion on it, I mean you can always study all these things. The other hypothesis is testing an issue. But keep in mind that when we are discussing about evaluations, we also should consider the hypothesis the formulation of prediction because based on that the analysis proceeds. Now here for this analysis part we have two designs the one is the correlational design often we can call it a bit of a non-interventional and non-experimental design because here we are not manipulating the variables.

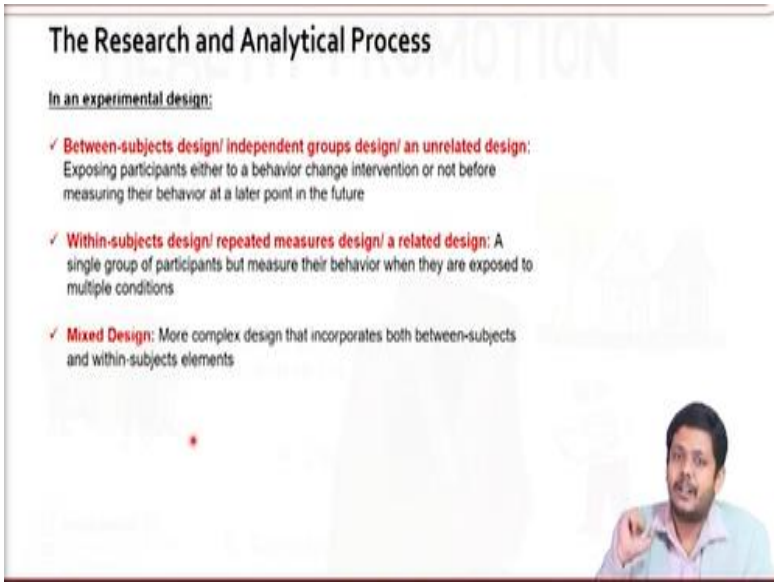
See here I have mentioned manipulating the variables. When we; are manipulating the variables in study that is the experimental design. Because see when we are I mean when we are studying health behaviour and we are studying the health behaviour in two different groups. And we are not doing any behaviour change intervention we have identified that there are certain behavioural factors or there are certain behavioural practices in one group and that practice is not present in the other group, a simple comparative study.

In that case we are just observing it. What we are doing is we are just only measuring the variables. So, what somebody intends to do its intention and what actually they do like this way and we have a comparison between two groups. This is called a correlational design or non-experimental design because we are not manipulating the variable but very important. When we need to manipulate the variables how do we manipulate?

Whether somebody is allocated to a behaviour intervention group or given the control condition. When I am doing that that smoking cessation app thing, I have given that app to some people those are my intervention group and I have put some people in the control group. So, in that case what I am doing is I am manipulating the variable. Here the variable is who is getting the intervention is we are not getting the intervention.

That I as a researcher I am doing it like so in this case this is an experiment. This is not a simple observation this is an experiment. So, here it is an experimental design this is not the simple correlational design.

(Refer Slide Time: 06:49)



So, that now we have understood. In health behaviour change interventions when you are doing the interventions you are mostly relating yourself with the experimental designs. So, when experimental designs you have to consider certain issues. You have the between subjects designs which is also called the independent groups design and unrelated group design and you have the within subjects' design or repeated measures design or related or I mean a related design.

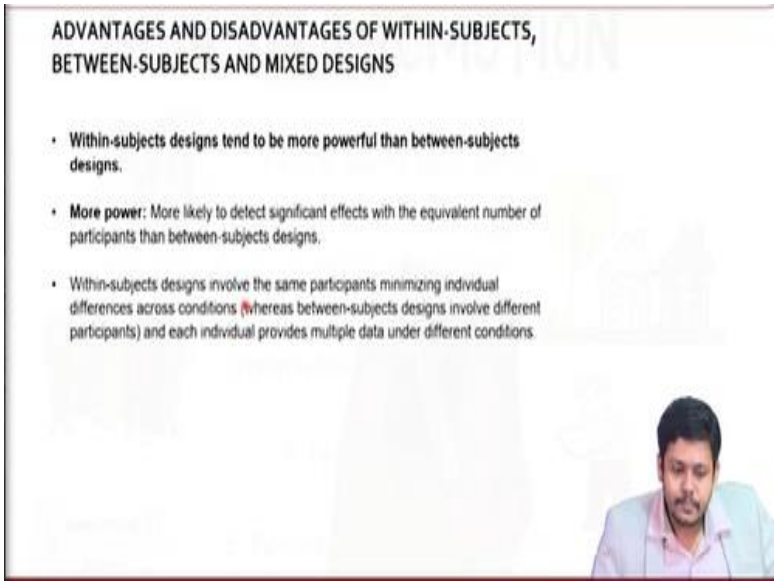
And then the third design is the mixed design which is in fact combination of these two designs. So, what happens with this experimental design I mean it is interesting you know between subject designs so independent group designs means you have two groups, different groups different individuals are placed within those two groups and you are studying them. You are giving intervention to one group and you are not giving intervention to another group.

That means you are putting another group and under the control conditions. So, this is you between subjects design independent groups two groups for example or the unrelated design. Because the persons who are getting intervention the persons who are not getting interventions. These two people they are unrelated they are not the same person. But when it happens within the same subjects a classical single group pre post design you are giving the intervention to the same people whom you have already surveyed.

You are not differentiating them in the two different groups like what you call it intervention or a control group, you are not differentiating them you are giving intervention to all of them. Now that becomes a single group design or a within subjects' design because here you are experimenting with the change within a subject. In the previous one we were experimenting with the changes between the subjects and in this within subjects design you are also getting the repeated measures reason.

Because you are measuring the same individual repeatedly it is not that the between subject systems you are not going to measure the same individual repeatedly. But here the difference that you are going to understand is through the repeated measures within the same subject that is why it is called a repeated measures design. And it is a related design because the same individual is being intervened and so the pre-intervention data and the post intervention data, they are related to each other like. So, combination of these two this gives a mixed design.

(Refer Slide Time: 09:23)

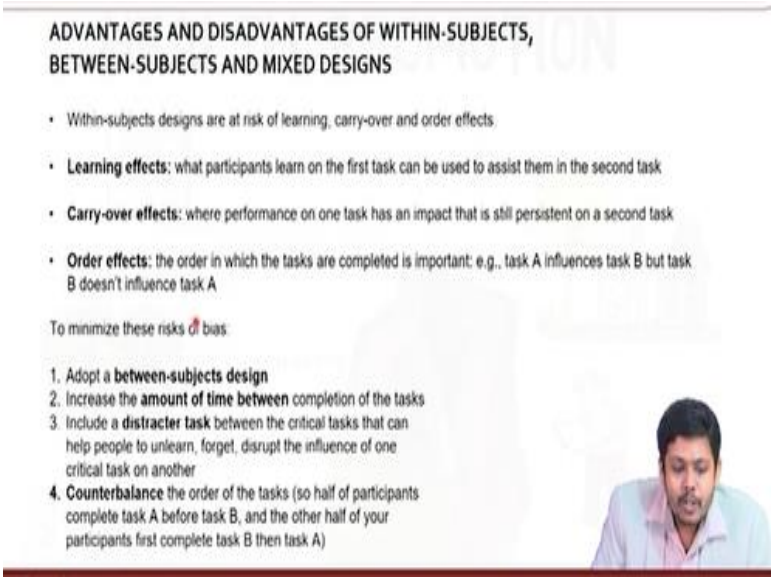


So, now that we have understood what are the different experimental designs. Let us discuss briefly what are the different advantages and disadvantages of within subjects and this between subjects and this mixed design because there are three designs you know. So, within subject's designs tend to be more powerful than between subjects design. Why more power? Because they are more likely to detect the significant effects with the equivalent number of participants as compared to the between subjects design or in fact even in the mixed design.

So, the most powerful designs are the within subjects design. What happens with this because the same individual you are not facing the confound or the inter individual level variables. The comparison group if you have a comparison group then you have to face the issue of confounding. So, here you are not facing that issue that is why it is giving you an extra edge that extra power the statistical power you are gaining through utilization of this design.

Now within subjects design involve the same participants minimizing the individual differences across conditions that I have just mentioned.

(Refer Slide Time: 10:36)



Now let us consider certain important effects. These are very interesting you know because the within subject designs these are I mean though they are having more power but there may be certain difficult biases that come into play. Some of the biases I mean we cannot discuss all the biases but three important biases in the context of health promotion education intervention are here are the **learning effects, the carry over effects and order effects**.

See classically when we try to understand the effect of a health intervention with in a between subject group what we do we give health intervention to a group and we do not give that intervention to another group and we finally differentiate or we finally see the difference

between the changes and we find out an effect size like this we do it. But in within subject group in a same individual when we are testing the same individual what we do is.

We let us first say for first six months we consider we do not give the intervention and we measure the person at the base line and after six months and then after that observation we give the intervention and then observe after the intervention suppose say after another six months. So, this is how we do it or we may consider giving the intervention at the first then after six months we withdraw the intervention, we do not give the intervention and after six months we again see.

And that is how the same individual will act as his or her own control that is how we do it in the within subject. But this methodology will incur what is called the learning effect, what participants learn on the first task. See when we are giving the intervention that is the first task the participant is doing and can be used to assist them in the second task. For example, if we are you withdrawing the intervention itself a simple intervention may be withdrawal of that intervention may itself be the second task.

But that is dictated by the first task that is the intervention proper. Suppose how to quit the quit smoking the proper app that is already there. The materials are already there now after six months I am not going to intervene on that part. But still the person who has knowledge of how to do it and what not to do it so he may go on recapitulating his learning. So, by that recapitulation what happens is you are not getting the pure effect of control.

What you are getting is the control with learning effect. This happens because a person has now learned what I mean how to quit smoking. Though you are not giving that intervention but that intervention is already there in that person's memory this is how learning effect occurs. Then there is the carry over effect, what happens is where performance on one task has an impact that is still persistent on a task I mean on another task.

Consider two equivalent I mean two related tasks like for example if you consider wearing mask and using sanitizers or using sanitizers and continuous hand washing. Now these are the two tasks. What happens is you carry over your learning you carry over your understanding of

performing that task to the next task. So, this is a derivative of the learning effect but this is called a carry over because your carry over performance on one task has an impact on the other.

Because your performance on hand washing only may ultimately impact your behaviour of sanitizer use. It may be like this way that a person who is hand washing frequently the person may think that okay I am hand washing frequently so I do not need to use the sanitizers. This may have a negative impact or the person may think that I am hand washing frequently and it is also a method of cleaning my hand so I will watch the; I will use sanitizers more frequently. This is a positive impact but the carry over effect may be. There the carry over effect may be negative one and maybe a positive one but it is there. So, we should recognize it then there are the other effects. The order in which we do it do we do hand washing first then after some time. Suppose use of sanitizers or do we first use sanitizers then after some time use hand wash like this. So, what they say is the order in which tasks are completed is important as for example a task A influences task B but task B does not influence task A.

For example, a person may consider that when I am washing my hands with soap and water, I do not need to I mean the use of sanitizers become implicit or it is I mean I cannot do it. But that same person may consider even though I am using sanitizers I must go on to wash hands. So, what happens here this washing sanitizer I mean washing hands with sanitizers are using sanitizers it is influencing the next part or it is impacting the next part that is going to wash hands.

You see in this context the carryover effect and the order effect become bit related. But what you have to remember is you have to differentiate between these two effects in this way that carry over effect one task is having impact on the other task. Now one task and the other task suppose task A and task B they may not be ordered, they may be mutually disjoint tasks. But in order effect what happens is task A and task B their order either task A is happening first then task B or task B is happening first in task A.

So, in this case when it is ordered you have to have this order effect in mind. So, how do we get over this effect? We have to minimize this effect we have a good intervention at our hands. So,

how do we do that; we adopt a between subjects design to group design. Because for a single group there may be this learning effect, there may be this carry over effect and more obviously there may be this order effect.

Because the same group is performing the behaviours in two different orders and the two different orders the activities may have influence on each other. Similarly, for carry over effect you can easily understand but in different subject, different in between subjects group or two group design or three group design whatever it may be the same person is not keep or keeping on doing all these things.

So, you can easily get over with this carry over effect or order effects you can easily null and void all those things. It also I mean by you can also take care of these biases by increasing the amount of time between completion of these tasks through time you are you to using or you are allowing wash out. Wash out of an intervention is very important to consider all these things but for that you are giving extra time.

However, washout the context of washout in health promotion education HPE intervention is very intriguing. Because the duration of washout is there is no typical guideline for that. You have to find it out through your own research anyways but increasing the amount of time between the activities, it is another method. You can give a distracted task; distracted task means the task through which the person is distracted.

Suppose you are doing this task A and task B this is your idea but you know that task A will eventually influence task B. In order to remove that influence you introduce another task C over here. So, through that task C which is unrelated to task A and related to task B and is also unrelated to the outcome this is very important. Because if task C is related to the outcome, then what happens is then you have a three intervention basically.

Another intervention which you did not try to study but it is already there. So, whatever be the distracted task that should not be related to your outcome. Because of this this distractor task the influence of this task A on task B is now hampered. So, this is how you can do I mean the

activity of unlearning forget or disrupt the influence of one critical task on another. This is particularly important in case of order effects you can understand.

And there is this important intriguing I must say concept of counter balance the order of task. So, I mean counter balancing the order of task is requires it ideally deviates from the single group design. So, what happens with this? Half of the participants complete task A before the task B and the other half complete task B then task A. So, these were the two orders you allowed these two orders to different groups.

So, essentially in order to have the counter balancing effect through which you want to counter balance the order effect you have to employ a between subjects design.

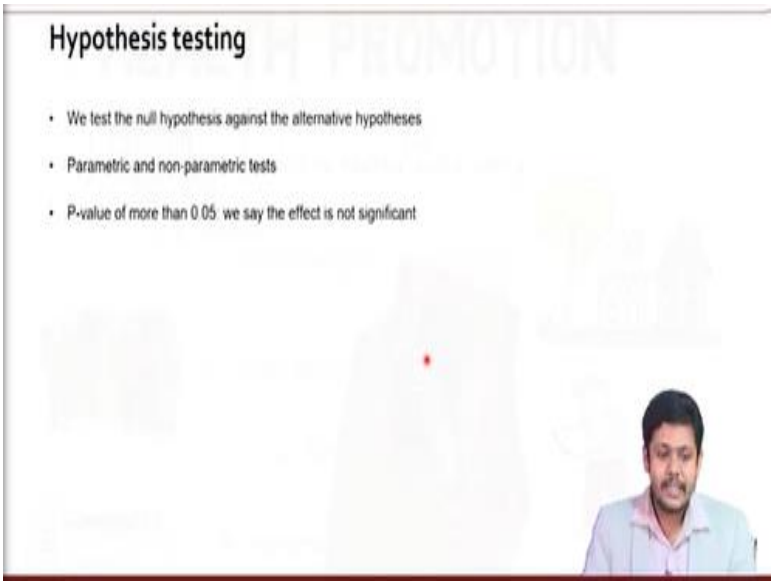
(Refer Slide Time: 20:06)

Then there is the issue of mixed designs. The importance of mixed designs is that they allow the experimenter to assess the impact of order effects. In the same group you can have the two different orders and also you can have a control group which is following a single order. So, when we are differentiating the outcome effects when we are finding the effect size what essentially happens is you have the order effect already there in your single group design and also in your control group you have the true actual effect.

So, the difference between these two will give you, not only the effect of your intervention A and B through in between analysis. But through this comparison single group and the control group comparison this will give you the remaining effect size. So, this is again important. So, the basic concern with mixed and between group designs is because there are certain other socio-demographic and cultural factors which may act as your confounding factors they require more participants than the single group design.

But as you can understand that with the help of these different designs you cannot only study the effects properly effects of interventions properly you can also study the effects of different biases that is very important.

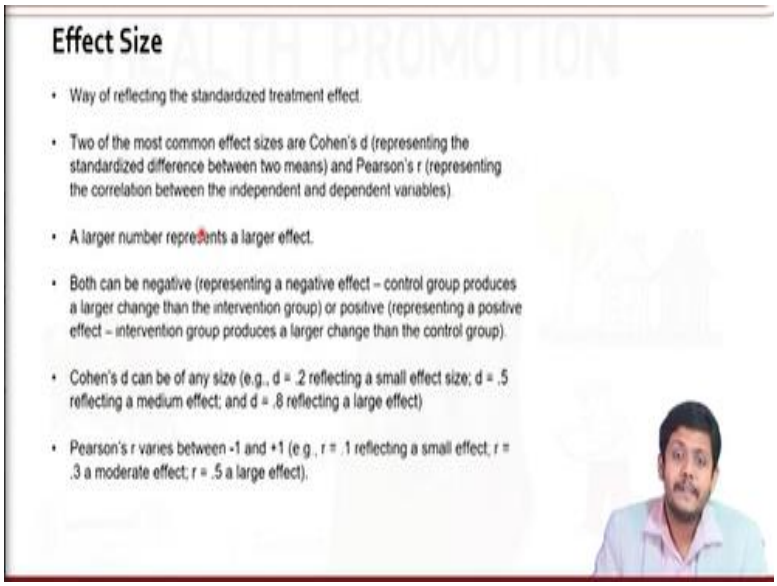
(Refer Slide Time: 21:28)



So, this was hypothesis testing that I started with and there are different non-parametric and parametric tests. I am not going to go into details of this because this is not the essence of this lecture, you must understand that these tests are there this is how you are going to analyse. But the whole analysis part is whole another domain and probably in another course we can discuss that. So, what we remember? We remember a p value cut off is 0.05.

So, whenever we have this cut off if it is more than that and for a test if we have more than that then we can say there is no difference we take the null hypothesis. But if the p value is less than 0.05, we accept I mean we basically say that the there is significant difference.

(Refer Slide Time: 22:17)



But hypothesis testing p value use is dependent on sample size. But in effect size analysis we do not have that issue of sample size. What happens with effect size it, it basically reflects the standardized treatment effect by treatment effect we mean the intervention effect treatment effect is a standard terminology used in trials. But in a whole effect size represent the intervention effects. More the common two effect size what you can have or what we can understand one is Cohen's d is for difference pre post difference.

And the difference between those two differences you have d 1 for group one and you have d 2 difference for group two, and the d 1 - d 2 this will ultimately give you your Cohen's d and you can have the effect size you can represent them with numbers. The larger the number it represents higher the effect size and also Cohen's remember I mean they can this Cohen's d and the Pearson's 'r'… r means it is correlational coefficient.

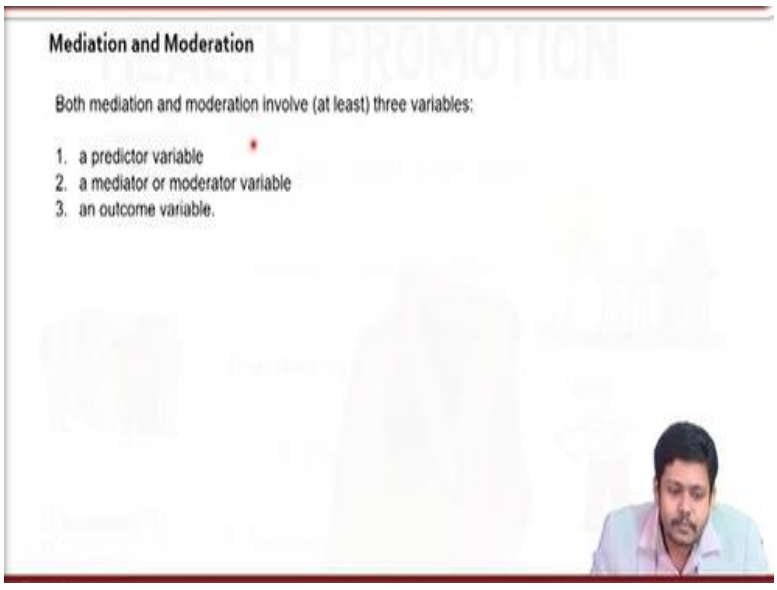
They can be negative also negative means a negative relation and positive means a positive relation. The higher the value stronger the relationship.

(Refer Slide Time: 23:31)

So, this is how you can use effect sizes as I was discussing p value is more related with the sample size but effect sizes are usually independent of the sample size. Effect size mean in terms of Cohen's d and I mean the Pearson's 'r'…. whatever you can say these kinds of effects what we measure through difference in difference approach I mean these are basically independent of the sample size.

So, for a very small but significant you know the important word is you have to ascertain the sample size through appropriate calculations then only the small sample size or lower power of a sample that can even yield the similar effect size. But the p value may ultimately different.

(Refer Slide Time: 24:14)

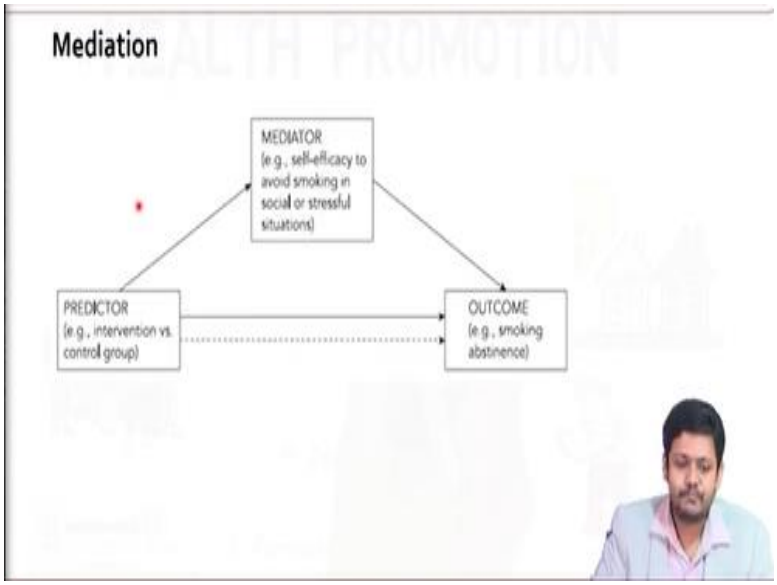


Now we come to the basically the last part of this analysis segment. We come to mediation and moderation effect. So, remember mediation and moderation they basically include a third variable in the model. So, we have task A and task B or now let us consider only task A, task A is your input. You are putting you are giving the task A to lead the outcome; you are giving your task A suppose is use of mask and your outcome here is prevention from covid or not getting infected.

Here you have a third variable, the third variable let us consider is suppose hand washing or using your crok of elbow to cover your nose and mouth when you are coughing like these. These are the appropriate behaviours any one of them let us consider it as a third variable. What happens is when you are protected from covid19 illness it is not only the effect of our intervention variable that is use of mask there is some effect of this third variable.

Now this third variable we have to consider whether this is a mediation or a moderating variable. So, basically what happens what is the common characteristic of both these variables, it is a predictor I mean in the model there are three variables. The predictor variable, a mediator or a moderator variable and an outcome variable.

(Refer Slide Time: 25:50)

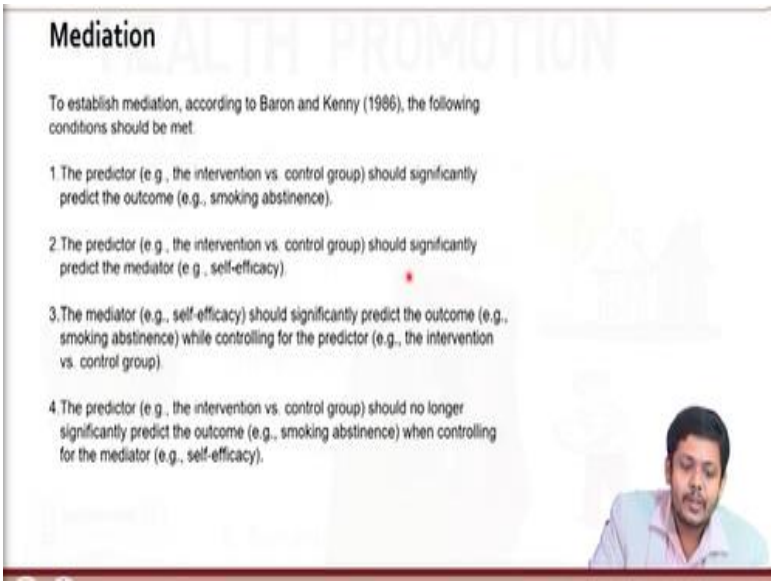


This is a depiction of a mediation process. You have a predictor; you have an outcome they may have a direct effect like interventions versus control group I mean giving interventions over here

is your predictors or the input variable and outcome is here suppose consider smoking abstinence. The mediator here is self-efficacy as we have considered in the SCT model. We are discussing in the previous lecture in SCT model self-efficacy is another mediating variable.

Why? Because the predictor the intervention it moves through this self-efficacy part and that self-efficacy ultimately leads to this smoking cessation.

(Refer Slide Time: 26:32)

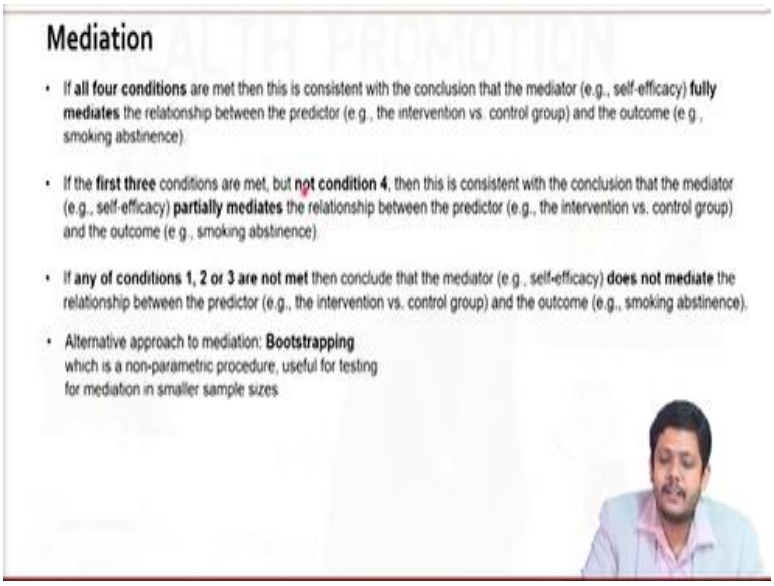


So, these are the criteria for mediation the criteria have been laid down by Baron and Kenny. They have worked extensively with mediation process and what they say the first one is a predictor should significantly predict the outcome. First is about this direct connection this should be there then the predictor should significantly predict the mediator. See predictor is significantly predicting the mediator.

So, this is significantly predicting the outcome direct connection is there and this is there. Then the mediator should significantly predict the outcome. So, the mediator is significantly predicting the outcome. So, this significant prediction this trio significant prediction this differentiates a mediator from a confounding variable. Then the fourth part is the predictor should no longer significantly predict the outcome when controlling for the mediator.

See this is important, when you are analysing the relationship between predictor and the outcome considering the mediator in your model then the flow should go through this only. The predictor is significantly connected to the mediator, mediator is significantly predicting the outcome. But in that model the predictor should not significantly predict the outcome this is very important, these are the four criteria.

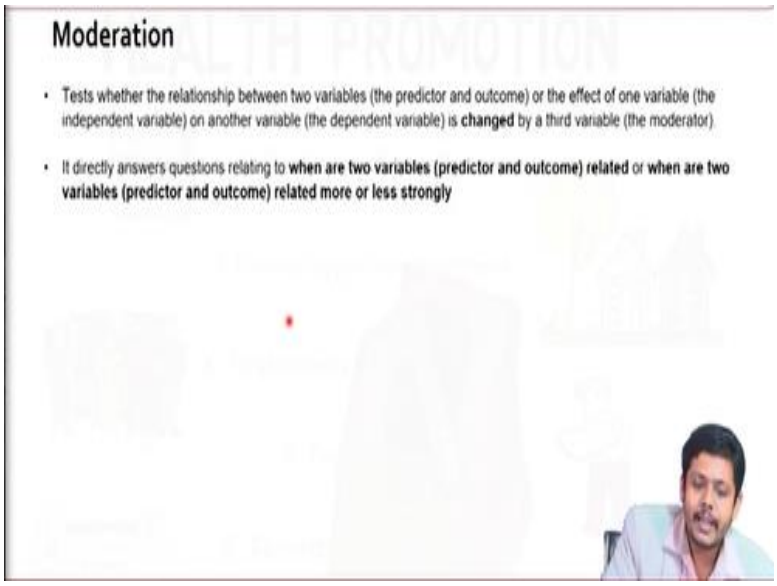
(Refer Slide Time: 27:53)



Now let us see what we call as a fully mediated effect, what is a partially mediated effect and what is no mediation. If all the four criteria are fulfilled that is fully mediated. If the first three conditions are made but this condition is not made that means this is significantly mediated connected this is significantly connected or the effect is the significant effect is there. But even in the presence of mediator the predictor is again significantly predicting the outcome.

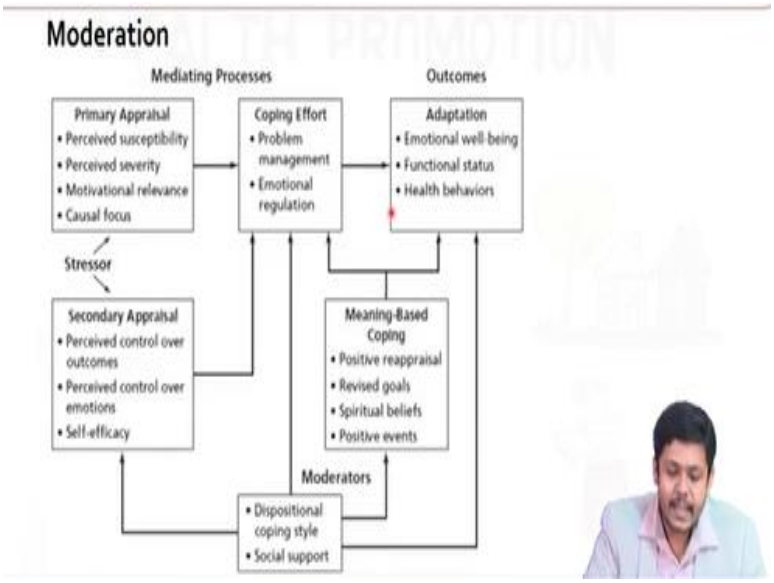
This is also there then we call it a partial mediation. But if 1, 2, 3 are not met then we say there is no mediation at all. Apart from this typical technique of Baron and Kenny we may also propose another technique that is called bootstrapping which is also a non-parametric technique. It is I mean it is kind of a statistically rigorous technique but it is useful for testing mediation in smaller samples. Because see in this case if we; go for a path analysis approach also it will require a larger number of samples.

(Refer Slide Time: 28:56)



Now let us come to moderation. So, what happens with moderations? It directly answers the questions relating to when are two variables that is a predictor and the outcome variables they are related or when are two variables that is a predicted and outcome variable so more or less strongly. That means the moderator it in fact changes the strength of association that answer is given by analysis of moderation.

(Refer Slide Time: 29:20)



Here also you have three at least three variables. Now you remember seeing this diagram this is taken again from Glanz's textbook in the previous lecture on health behaviour models here you can see these are the moderators. So, these are the moderators, why? Because the outcome is

happening here. See the outcome is happening here and I mean what do you say the coping style social support and from all these I mean the outcome is taking place.

I mean the path you should follow this is happening. So, from stressor if we consider from stressor to outcome, these are all the mediating process, these are all the mediating variables. Because the primary appraisal is there then the coping effort is there and then the adaptation or adaptation is in fact the outcome, adaptation of stressors, adaptation of behaviour from this effect of stressors this is happening. So, this is how these are the mediating process.

These are the process these are the variables through which the outcome is connected to the predictor what you can say is connected to the outcome. But these are the moderators it these in between phrases. Consider meaning based coping this is a moderator moderating variable because this is changing the effect of say how this coping effort is leading to outcome. Because here if we; consider this part in isolation you have this as an input and an output variable kind of relationship.

So, the relationship between the input and the output variable the strength of association it is changed by this variable itself. The meaning based coping thing itself, coping effort is related to adaptation in some way but that relation is changed or the strength of association is changed by meaning based coping in that model. But this is not mediating factor because this is the mediation criteria is not fulfilled by meaning based coping.

Because see meaning based coping can be I mean related to adaptation but consider coping effort and meaning based coping these cannot be causally related or significantly they may be correlated. So, in this case meaning by coping acts as a moderating factor because it is changing the strength of association. But it is not a mediator because it is not within that path of outcome to process.

(Refer Slide Time: 31:58)

Now we come to handling of missing data. Because we will be having certain missing data when we consider this health behaviour models and interventions, they may be due to data collection errors or they may be due to non-response or partial non-response of your participants. So, how do we handle missing data? We leave it as it did and we live with it and you analyse your data on the basis of the responses that you have and you may consider excluding the missing data part and analysing the remaining part.

But also, you can have an ITT approach the intention to treat approach in analysing health behaving intervention data with missing data component also. How do you do that? You just impute the variables or you consider the last measured data for that person. Suppose in observing base line observation you have some data and in observations suppose in a second follow-up observation you do not have any data for that participant.

You can consider the data from the first observer first follow-up observation for the second follow-up observation to fill up that gap of missing data and you can also have multiple imputation techniques also.

(Refer Slide Time: 33:05)

So, this brings us to the end of this lecture on analysis of health behaviour change data. What we have learnt is there are different designs like correlational designs and experimental designs and under experimental designs you have your between subject variation design you have your within subject variation design and they may be combined into a mixed design and also what we have learned is the mediation and moderation should be considered for contextuality as a part of this third variable.

Also keep in mind the thumb rule in handling missing data you should consider ITT approach by utilizing the previous response or by doing the multiple imputation technique.

(Refer Slide Time: 33:47)

So, this is bit the analysis part is a rigorous I mean methodologically rigorous part. So, I recommend you go through this a bit in detail and understand the concepts these are the resources as we were discussing in many other resources and in many other lectures. These are always your resources for this discussion. Thank you.