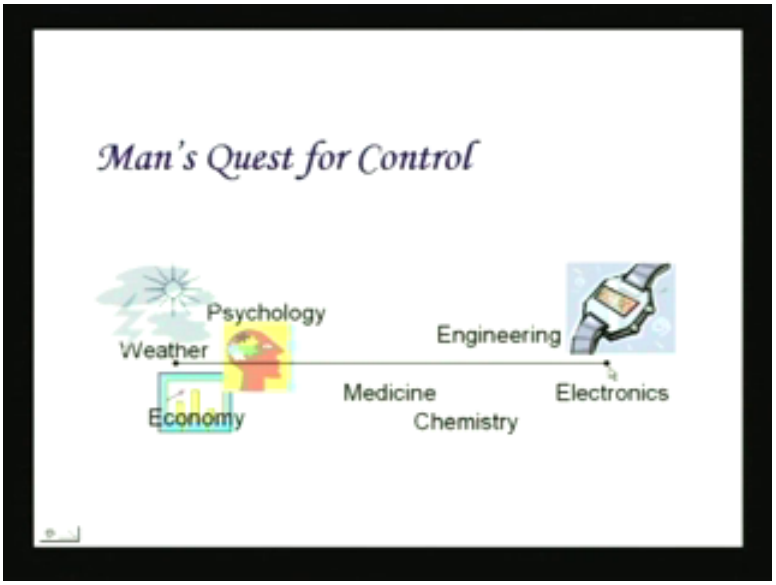
Six Sigma
Prof. Dr. T. P. Bagchi
Department of Management
Indian Institute of Technology, Kharagpur

Module No. # 01
Lecture No. # 26
Design of Experiments – Overview

Good afternoon, we resume our lecture on Six Sigma and the topic that, I am going to be discussing today is a very key one in the process of achieving parts per million types of defects this technique, is called Design of Experiments and it is as I said it is integral of the six sigma methodology.

(Refer Slide Time: 00:42)



Let us take a look at, why is it that we need to do experiment and what is it is role in the full range of quality engineering, which is actually today it has culminated into six sigma that is the state, where we are today. If you look at the picture today, you can see the clouds there and underneath is weather then of course, on this side we have got psychology and I have got the economy here doing its ups and downs, on this continue and if you come to the other extreme you got this digital watch, if I may show you my digital watch this is at other extreme of this knowledge that is there.

On one end we have got very little knowledge about say something like the weather system, and on the other hand when I when I talk about digital watches, we are right there, we know 100 percent about how digital watch functions. In between our all the different system that we have around us medicine for example, chemistry, engineering, electronics on there all over the place they are on this spectrum.

Now, what is the difference between systems on this side which is to my left, and the the systems which I have to my right hand side; it turns out that man has always tried to control systems, whatever we have around us, we always made an effort to try control it. And one of the things we have tried to do, is to try to first do basic research and of course, we began with observation and then it seem like some theory was behind it, then we put in our a lot of effort to try to understand the basic driving forces behind that process. In the process we ended up with some fantastic theories, the result is something like the digital watch or the computer and the today's communication device and so on.

Their knowledge has progressed to the point, we can write the exact equation how a digital watch functions for example, this is not so if you look at the weather system. The weather system has too many variables of you remember my cause and effect diagram there are so many factors, there are multiple factors that affect the weather system, and it turns out they are very very little is very known how the weather system behaves. So, we cannot certainly write down today's, as up today's state of knowledge we cannot write write down equations, exact equations to describe how temperature is going to vary, how humidity is going to vary and so on so forth, that we are not able to do today, as far as weather systems are concerned.

Same thing goes for psychology, we are not able to write down the equations and also the same thing pretty well comes pretty close to economy when it come when it comes to talking about the economy, we know very little about the real process that is there and of course, theories are there to try to sort of you know and those are based on primarily observations.

In between of course, there are all these other systems, but we people we men and women, we cannot really live with a system that is highly unpredictable, and we had the

we are at the mercy of clouds and sunshine and so on so forth; we really cannot do anything about them, we really cannot with the system like that. So, there has been a constant effort to try to take systems which are on this slide, on this spectrum to try to gradually push them to the right hand side, so we can control it.

This has been our continuous question, in fact it began probably 1000s of years ago and today, we have some knowledge in medicine, we have a fair amount of knowledge in chemistry, we have a lot of knowledge in engineering and certainly in areas like electronics we have really got all the equations known and determined and so on. So, basic result research and also applied research it has advanced to the point on the right hand side that we know everything about it, and we can write down the equation.

Now, if you are trying to if we if our desire is to take systems that we know very little about and we want to push them to the right hand side, that we can control them, how do we go about doing it, well we can just stand around and observe that is like one way, just observe and try to see there is any kind of correlation between, what we see. And what happens perhaps there is some driving, some some cause in effects relationships that is how, perhaps these complex phenomenons take place for example, rain and so on and so forth.

Then of course, the process can be further modified and this is where we have tried to use our intelligence, what we have said is suppose we are able to identify some factors and we are able to construct, what we what we call the fish pond of the cause and effect diagram; and then we manipulate some of these factors and see the effect. So, what we are really doing here is, we are intervening, we are intervening into the process, we first try to figure out what are the different factors that might be affecting a particular process, then we manipulate these factors in a manner that will produce hopefully some effect there, now this is the experimental approach this the empirical approach.

You have manipulate some of the input conditions you take a look at the output, if can see some sort of dependency, there some sort of relationship then of course, you got some clue as to how this process this prestigious black box, how it operates, that is something that we can really figure out. If it, this is what is done a lot, if you look at

engineering, if you look at chemistry, if you look at lot of these things, a lot of knowledge in this area, has been basically acquired through empirical work, through experimental work. Now, if we did experimental work that is fine, but it turns out many factors are affected by multiple factors, many systems are affected by multiple factors.

That means; somehow I have to basically take four, five factors and I have to manipulate them together, in some complex way this is not a easy task, it is not very easy to manipulate four, five factors together then see the effect, then come back say come back and say that the response, a particular response was due to factor a, another response was due to factor b, another response was due to factor c.

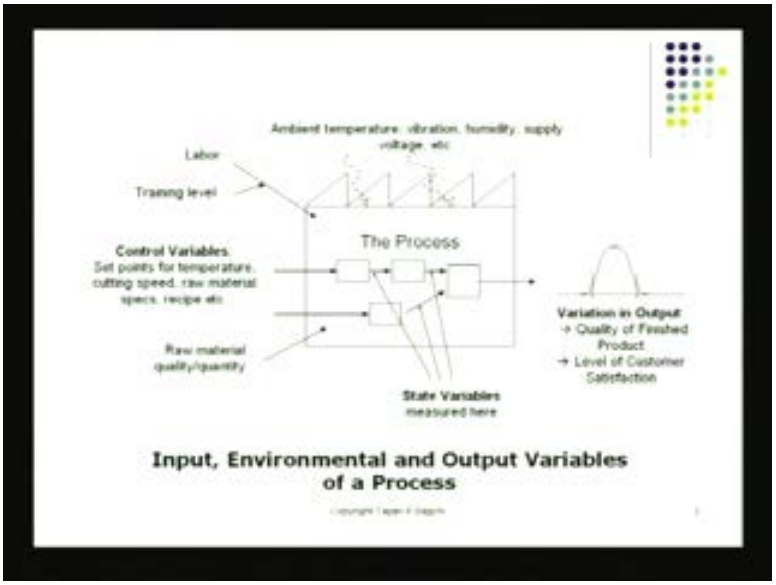
And perhaps something was also going on in the background when factor d was acting, this is not very easy to do, if you just do one factor at a time trials I will show that later; I will show why it so difficult to do one factor at a time experiments and hope to see that you get understanding of the process. So, what we needed here was a technique that could take multiple factors together and give us a procedure give us a methodology that could again give us insight as far the factor effects were concerned and indeed if there was any interaction between the effect, we should also be able to uncover that.

This was done around world war two, and the gentleman who did this his name is Ronald Fisher, sir Ronald Fisher in UK, he basically come came up with the scheme and he did some mathematical work also, and he he was able to find a technique which is later on in mathematics it is called, in statistics it is called ANOVA, Analysis of Variants; this was the method by which you could find out, which of these four factors really has an effect, indeed if two of those factors interact if that is significant this is also you can find out. And also this led to slowly and slowly gradually getting to an equation an empirical equation form then of course, other people got involved and they tried to evolve the basic understanding of process later on as time went by.

So, this technique this technique of manipulating multiple factors this came from this professor Ronald Fisher, we will see how this was done that is like the cracks of it. So, in fact what we are talking about is, we are talking about multiple factors manipulated together and some methodology that will let you see, the effect of each of those factors

and perhaps their interaction also; this is the sophisticated empirical approach, this is very different and much more advanced when you compare that to one factor at a time trails, it is quite different from that, let us see how you get into this.

(Refer Slide Time: 08:37)



If you look at any of the today's systems if you look at any of the today's systems, you do not have to really look at the digital watch that is far to complicate, if you look at of factory for example or a machine you know something that is doing grinding or something that is doing welding and so on so forth.

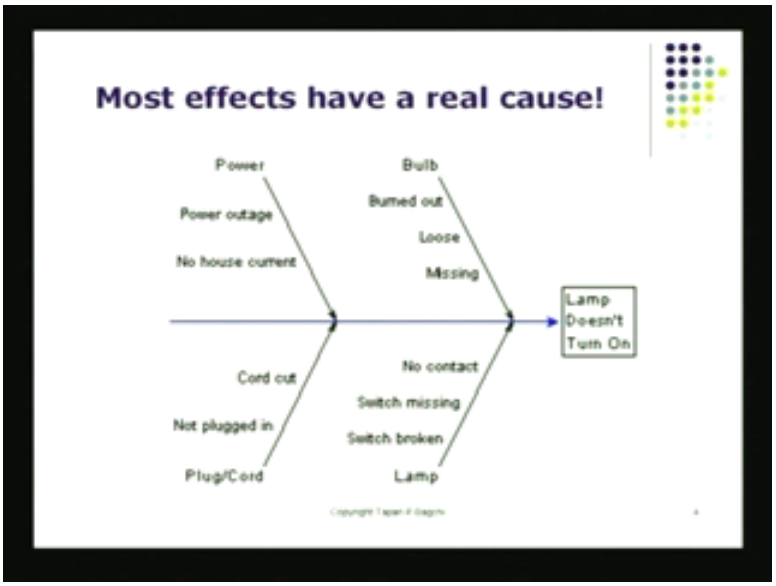
If you look at any of these systems generally speaking, what you will find as you see see on the slide, there are multiple factors that might be effecting the process, some of these these we call control variables and these are of course, on the input side I have got raw material, I have got labour, I have got training level and so on so forth. Then I have got process process conditions, various process conditions which are there, these are going to be influenced by the control variables, these are going to be factors that might be in impacting the process that is inside.

In addition of course, we have also got ambient temperature vibration humidity and so on so forth, these are external conditions and these conditions are not necessarily always in

our control, and somehow we have run the process in presence of all these noise factors. Anything, that you are not controlling we regard we the professional we regard those as noise factors, so me how what I have to do is I have to figure out, what is going on inside this black box which is the process.

If I had the equation of course, there will not be much difficulty in coming out and writing down the equation, the exact equation that will show the output as a function of the input factors. When that equation is not available I basically have to do empirical work, and this is done a lot in R and D labs, in the R and D labs they manipulate these different factors try to and they try to see, what is the effect of it.

(Refer Slide Time: 10:20)

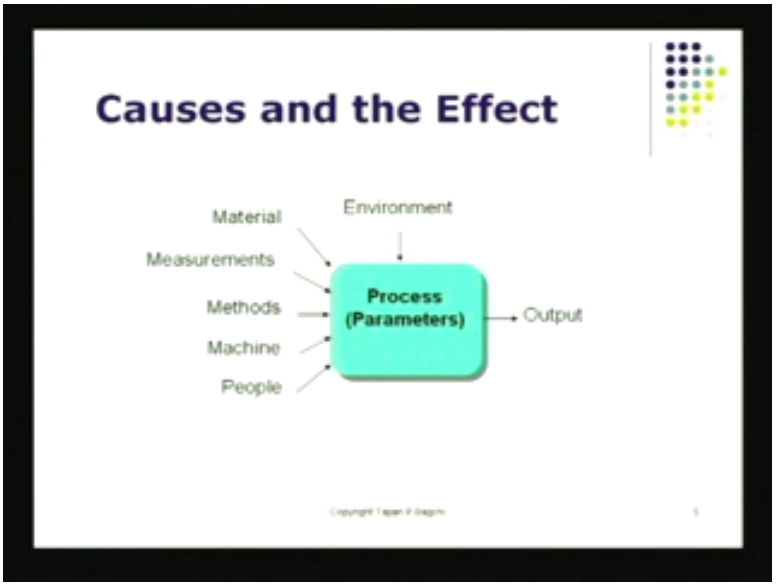


Let us see how we do that, the beginning of course, is again this fish pond diagram where I have got some speculation of how a particular process works and I have listed out the factors here, and the output is on the right hand side which is effect of all of these. Now, if I have to study these if I have I study these different factors I certainly need some special methodology, I cannot just do that the way ordinarily experiments are done, I cannot do that by manipulating one factor at a time just cannot do that, I have to manipulate multiple factors together.

Because, the real process has the involvement of all these different factors there therefore, what we have to do is we start with the cause and effect diagram and then we come up with a procedure which will allow me to set the different settings and these are called treatments. Of these different factors which are there, then look at the output and then do some special data analysis which will let me see, whether factor A was important it had an effect strong effect on the output or B was important or C was important or D was important or was there an interaction between A and B or A and C and so on so forth.

Now, the the process the experimental process involves all the factors together I am not doing a one factor at a time trail that is not the way I can move from left to the right, I cannot go from the state of zero knowledge to perfect knowledge, I cannot do that by manipulating one factor at time. Partly because, factor effects many times they interact and that is a very complex phenomenon and that is why I have got to manipulate multiple factors together, and this I precisely what is done in DOE, Design of Experiments is the special method that Ronald Fisher came up with.

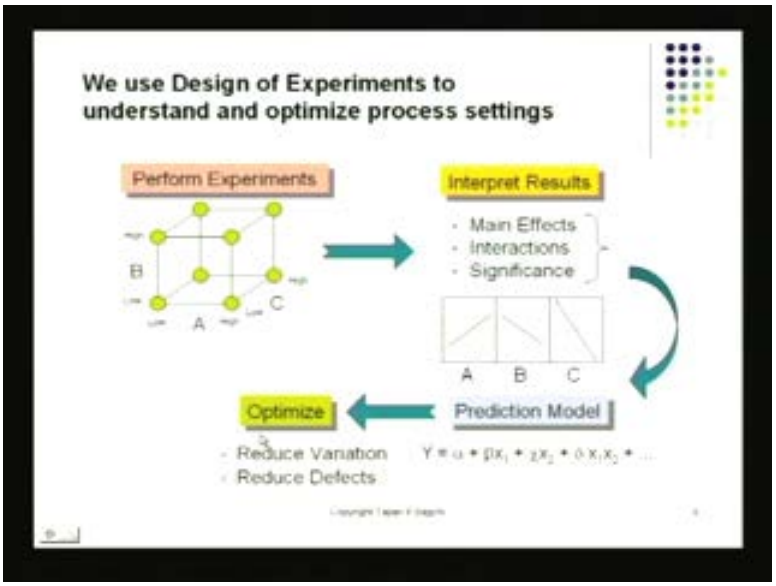
(Refer Slide Time: 11:59)



So, if you look at a process, I have process parameters inside and I have got all these different factors and I see this cause and effect dependency there, the effect is the output

and the causes are any of these factors which are to the left of this box there. And my goal is going to be to manipulate these factors in some systematic way, so that I can work out the dependency of the output to this input conditions; this is basically what we are trying to do, what we would like to able to do.

(Refer Slide Time: 12:27)



Now, let me tell you in a snapshot, let me tell you how this session is going to go and how the other sessions and design of it is, how they are going to go say you got three factors, you have three factors and notice a box there, notice a box to the left this is actually 3 D matrix, that is all it is. Now, when I am trying, when I am speculating about the effect of A, and the effect B, and the effect of C, what I do is I construct this matrix this matrix is an arrangement that led lets me manipulate simultaneously all three factors; in a particular way, in a special way and let me show you, let me actually show you, how this manipulation is done.

We will assume for a minute that factor A has two alternative positions, factor A can be set at low level and also it can be set at a high level, so in fact factor A has only two levels, for simplicity I am assuming that; factor B also let us say has two levels low and high, factor C also has two levels low and high. If I combine two times, two times, two I have got eight combinations, so there are eight different ways in which I can run this

process, when I am manipulating these factors A, B and C each can be set at two levels each can be set at two levels, what I am really trying now is, I am trying to sort of see, can I combine two and then two and then two, can I combine them and that combination is done by this special matrix there.

And I am going to tell you a lot more about these matrices, these are called DOE matrices, these are special factor manipulation plans, these are special plans that allow you, that really enable you to conduct these different experiments appropriately adjusting the different treatments that this process is subjected to.

What is the effect of all this in between you see that fat arrow there, transfer the data comes out, I run each of these experiment, so for example, this experiment here will be run at the setting up A at high, setting up B at low and setting up C also at low, so I have got high, low, low, that is going to be the setting for this experiment. It will have three factors set at high, low, low level, when I conduct the experiment, when I run the process by holding A at high, B at low, and C at low, I get some response, I record that response, that is like some of the outputs from my first round of experiment.

My second round could be this one, when I got A at high level, B at high level, and also C at high level; so the setting for this experiment setting for running this experiment is going to be high high high, I do the same thing for each of these corners there. And therefore, I end up running eight different rounds of the experiments, I have now run, because of this scheme there, because of this matrix scheme there I have run eight experiments.

So, the data that I have now on this big fat arrow there, this data now has the results of eight experiments, then there is a data analysis procedure, I apply some data analysis procedure to the data that comes out and the result of that is going to be, I will be able to discover calculate actually main effects these are going to be the single factor effect, effect of factor A, effect of factor B, effect of factor C. This I will be able find out individually, even if I would run the eight experiments by manipulating all three factors together, the data analysis procedure is so good, that it gives me separately the effect of A, the effect of B, the effect of C that I am able to see, these individual factor effects

these are called the main effect.

Now of course, there are other complications where the effect of factor A and factor B they may interact, this also you can find out from the same from the same data, the data that you have there from the eight trials, you will be able to find two factor interactions and perhaps even three factor interaction, because I have got A, B and C three factors, they may interact in a complex way, that also I will be able to uncover by doing this special data analysis technique.

So, in fact the reason is, these plans are special plans, these plan plans are actually very special plans, the scheme is a very special scheme that scheme is unique and that will allow me to separate the effect of A, effect B, and effect of C, and also the interaction of the two factors two factor two factor interactions, also one, three factor interaction. Because, there are three factors they will be one in fact it will be it will be just a setting of certain values of A and B and C and I will be see, the interaction effect of all three working together.

Now of course, there is some other line written there notice here, there is I have got a statement there called significance, I have got a phase there called significance, I have got main effects interaction and significance significance is a statistical significance, it is a statistical term that actually says the effects that I notice, they are going to be significant they are going to audible in the background of noise, it is sort of like this.

Let us say, if this studio had a lot of a a it had a noisy fan running and I was trying to do my lecturing in the presence of that noisy or the blower that was there and I can simulate that by saying somehow I will raise the noise level and either, I can raise the noise level and let you see what happens or I could reduce my voice to the extent that it becomes almost indistinguishable to the background noise that is there, and let me just try that.

So, if you notice now my voice got lost, the effect of my voice chord and breathing and everything it got lost in the background of noise, in fact if someone was sitting next to me even next to me he would have a pretty tough time to find out what I was saying, when I talk about factor effects. Each of these factor effects are significant relative to

noise, relative to background noise, because that background noise will be there, but what statisticians have done it, they do not say that the effect is significant unless it is like a multiple you know significantly higher than what background, whatever that background level of noises; that is actually how they figure out when it the effective significant, I need to worry about it or I can just let it go without doing anything.

Now, after I have got these effects calculated I can do some plotting and notice here I have got the effect of A plotted out, this like when A goes from low to high setting, this is what happens to the response, the response goes up when I move the setting of B from low to high the response comes down. And when I manipulate C from low to high again the response comes down, out of these factors A, B and C you can clearly see, that C has the largest effect, C has the strongest effect; and if my goal was to try to maximize response, I would have set A at high level, B at low level, and C at low level.

If I would pick this setting for A and this setting for B, and this setting for C the combine effect hopefully is going to be maximum, the maximum value for the response that is like an empirical optimization that manipulation actually has been taken place empirically. What else can I do with the data that I have there, there are techniques available with the help of which, I can take the data that came out of that green fat arrow, I can take the data and I can come back with an equation and look at the equation there Y is the response, I have got alpha plus beta times x 1 plus you know chi times x 2 plus delta times x 1, x 2 and so on so forth.

So, this is now a mathematical model and I can construct this model by regression analysis using the same data that I have got over here, I can use the same data I can bring it through and I can do multiple regressions and I can find out this equation, I can find this equation.

Now, because this equation has been derived using empirical work, we call this an empirical model, what the use this empirical model, remember the ultimate goal is for us to get in there and to make everything controllable like that digital watch, now fortunately for the digital watch, we did not have to conduct experiments. Here, we could study ohms law, Kirchhoff law and many other things, that that was provided to us by

people from physics the Ph.D from physics, they gave us the equations, and they gave us those electronic equations and so on so forth; and we were able to design this watch based on those equations.

In this instance, when I am trying to manipulate factors empirically I do not have those equations, but I can come up with the proxy equations and that proxy equation is this regression equation, which you see at the bottom of the slide here, Y equal to alpha plus beta times x 1 plus chi times x 2 plus delta times this things, this equation is the empirical equation, this empirical equation is almost as good as that original theoretical equation, if one able if one was able to construct it. What is the use of this, well I could go to the next state, I can optimize this function I can find the values of x 1, x 2 and so on so forth; that will maximize Y or minimize Y whatever be the be the desire that optimization, I can now do using this empirical equation.

So, just see where we started from and where we are, we started with the cause and effect diagram which is like this one, and I conceptualize them identified some of the factors and I used them in my experiments, and I constructed this matrix this is a special matrix and after I ran the experiments collected the data figured out the main main effects and the interaction effects, constructed the regression model and then of course, I apply optimization.

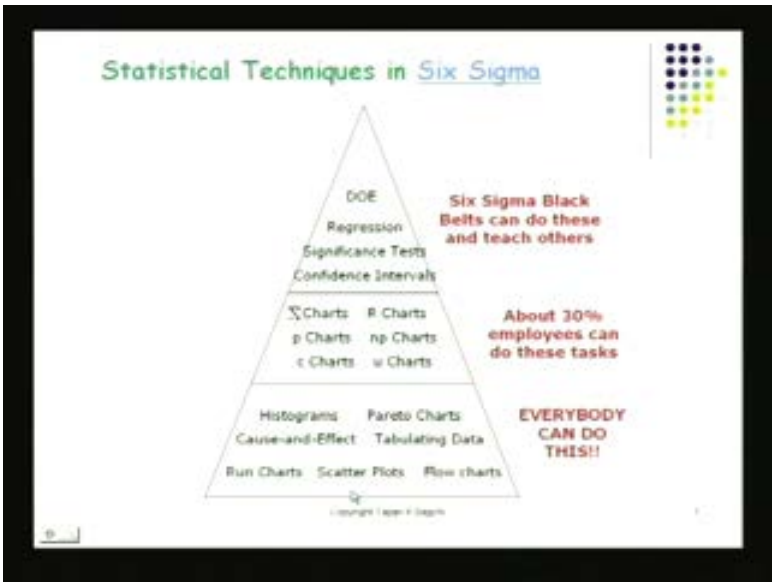
Now, what do you do, what what do you really try to optimize well, if it is like a product system and we try to maximize productivity, if it is a scheduling system I may try to minimize, the time to finish all the jobs, if it is a chemical process I may try to maximize utilization of raw materials, I may try to maximize the utilization of anything else that is there or I may try to maximize yield these things are possible when I apply optimization.

And this is quite easy when you got equation in front of you, whether it is based on basic theory or it is derived by empirical work like we done, in a lot of work in six sigma, we would be forced to derived actually these, construct these empirical models; and therefore, six sigma experiments also are going to be special experiments, they are all going to be guided by DOE the general principle of DOE, what is the effort, what is actually the aim of six sigma, six sigma tries to reduce variation in quality.

If you got any kind of variation in quality for example, if you look at the output of this process here, the output is highly variable and this may not be satisfactory for a customer, what we would like do is, we like to make this process as tight as we can make them and once we make the process tight it is going to satisfy the maximum number of customers. So, our goal is going to be then to try to shrink the spread of the process and I will show you later how we end up doing that, this is this is to be by us we are the specialists, we are going to be doing conduct some experiments, and if you follow the domain procedure of six sigma, define measure, analyze, improve and control, if you follow these this five step procedure, at the improve stage we are actually using DOE.

And that is when we are trying to manipulate the different variables, find the effects and also optimize the process all those things could be done, with the sole goal to try to minimize minimize variation in the process in the output of the process; that is like one goal also in the process we will end up producing defect that is like also another objective of conducting experiments.

(Refer Slide Time: 24:47)



There are many techniques available of course, many statistical techniques available and you already seen histograms, pareto charts, fish pond diagram you have seen tabulation of data you have seen run charts, scatter plots, flow charts, these are techniques that

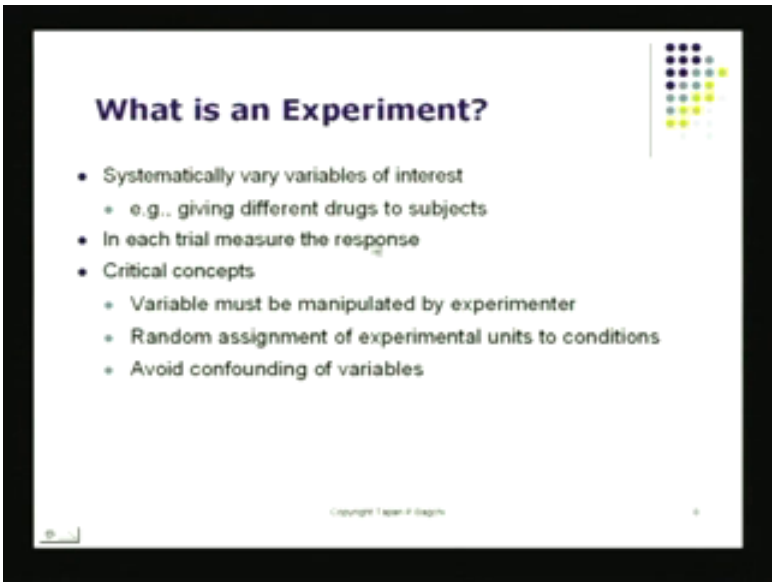
almost anyone can use. In fact, the the work force should be trained on these techniques for sure then of course, the engineering types of the production supervisor they should be able control the process online and they can use control charts.

So, training also should be provided at this level, so that they can manipulate the control charts, they can actually use the control chart to try to locate any kind of defect that is there for which there is an assignable cause. Then of course, we got these techniques and these are the most advanced statistical techniques, these are part and partial of six sigma black belt black belt training and the black belt people they can teach this to others also.

What are the techniques involved here, we already heard about DOE, we also heard about regression models, so there are certain other things called significance tests and confidence interval, these are advanced techniques which are all used if you are trying to optimize the process; and six sigma uses this the top of the pyramid six sigma very heavily uses these techniques in trying to do it.

So, every black belt should know techniques from top to bottom, every black belt should really know all the techniques that are there and specially they should really be focusing on this and they should be able to help, other people who are trying to use DOE or regression.

(Refer Slide Time: 26:21)



What exactly is an experiment, I gave you an overview, now what exactly is an experiment, is the process of systematically varying the variables of interest, these are the control factors about which you got some speculations, and the idea is to try to make sure, you are able to see the effect and you are able to you are able to do this with by manipulating some of the input conditions and by observing the output; that is an experiment, we do experiments all the time.

In fact, it is surprising how often we are conducting experiments and we are very mindful of it, we are actually indeed we learn so many things by doing hit and trial, that is also an experiment, we make some change in the input, we change the input and we look at the output, and we try to see is there a dependency there, if there is something then we learn something, we find out how this this particular system works. Every time we are doing a trial, we are measuring response, every time we are doing a trial in our experiment, we are measuring the response.

And some of the critical concepts which are involved in an experiments are variables are manipulated by the experimenter, random assignment of experimental units are there many many times what happens is, if you let us see you are conducting an experiment and you did some of the experiments early in the morning when temperature was low.

And the sun was not there, and then you conducted some other experiments at lunch time when temperature was thigh and the sun was out and the sun was pretty bright and so on, and then you conduct the same kinds of experiments you again conduct in the evening.

Now, you are manipulating only the variables in which you are interested, so you went to your fish pond diagram and you picked up your control variables from there and those are the ones you manipulate using the matrix that is there. Unfortunately while you are doing this, sun is also doing its own things, so it was cool in the morning it became warm then perhaps it cooled down again, that also might have an effect on the experiment, therefore, what you have to do is you have to randomize you have to randomize the order in which you run your trials, this tends to fool factors like sunshine and so on.

You will end up finding some effect of that, but it will get randomized they found the average there those things are going to cancel out, and this I am going to discuss again as we get deeper into this we are going to be doing this. And something else, we should also try to do and this is kind of a somewhat of an advance concept, we should try to avoid confounding of variables that means, in places we are manipulating one particular variable, another factor should not be confounding the same the same effect it should not actually be such.

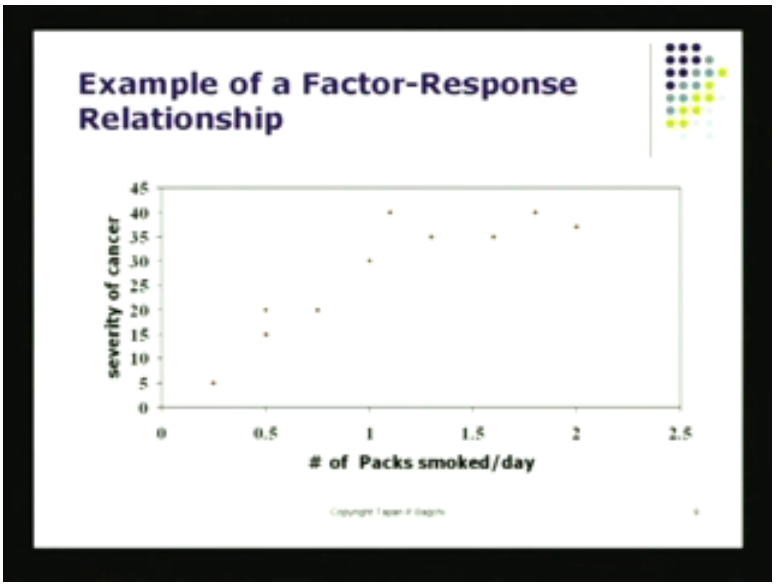
Then I am trying for example, to find the effect of one factor let us say that is my left hand and I am trying to do something with my left hand at the same time I should not also bring my right hand and try to do this, because then when the effect takes place I would not know is it due to my left hand, is it due to my right hand, here the effect of the my left hand and my right hand they can found it.

This makes it very difficult to figure out which factor actually was it my left hand or the right hand that produced the effect, the same the same idea we use also when we are designing experiments, we try to avoid confounding we let the variables stand on its own whenever we do settings of an experiment, we make sure if we are if you are manipulating a particular variable in a certain order.

There should not be another factor that is also manipulating in the same order, if you did

that if you did for example, temperature high and low and also the same setting we did the pressure high and low simultaneously we did not run any other trial, it will be very difficult to figure out what is what the the effect that is observed is that the due to the effect of temperature or is this the effect due to pressure, is effect due to pressure.

(Refer Slide Time: 30:08)



Now, if you conduct if you conduct some experiments and some very simple one of course, are there for example, this is of course, not an experiment, but you pretend that some factor is being manipulated and you are looking at the response so for example, number of packs of a cigarette smoke per day, so half of pack or one pack per day, one and a half pack per day, two packs per day, two and a half packs per day.

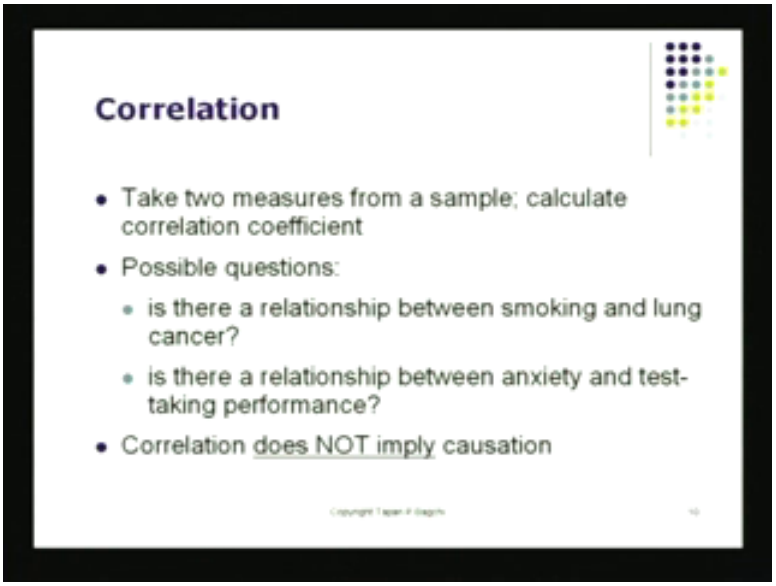
Now, something that is really correlated is well established a medicine is the severity of cancer, severity of people having cancer on some scale that is got a correlation with the number of packs of cigarette smoked, you see that raising trend there that raising trend actually says that there is a relationship between, there is a dependency between number of packs of cigarette smoke per day and the the severity of cancer that the man, that the person contracts.

So, this is actually is something fairly simple and easy to see, when I have got only one

factor, notice something else there are these, this not quite a straight curve, this not quite a straight line or a smooth curve there are some fluctuation, there are some random variations there and that is because, at some point in time the total effect of causing cancer that is not the effect of smoking only, it is got genetic orientation and lot of other factors which are there, that will differ people to people it may differ from city to city or situation to situation those also have their effect and those those who actually make people cancer prone.

Let us see, if I will if I will living next to a refinery and the refinery is being out all this sulfide dioxide and nitrous oxide and so on so forth, those are also have an impact on the person therefore, it is not just going to be his smoking habit, it is also going to be effect of this environment that is going to let him be kind of more like you to getting to cause to catch cancer. What we have to do here is therefore, we got to worry about combining not just using, just not one factor in my experiment, but also incorporating other factors also, these are those are that also speculated to be effecting the process, that also would be there.

(Refer Slide Time: 32:28)



So, correlation is something that is quite easy to see for us for example, we see correlation here and this is just defined here, that correlation is something that we find to

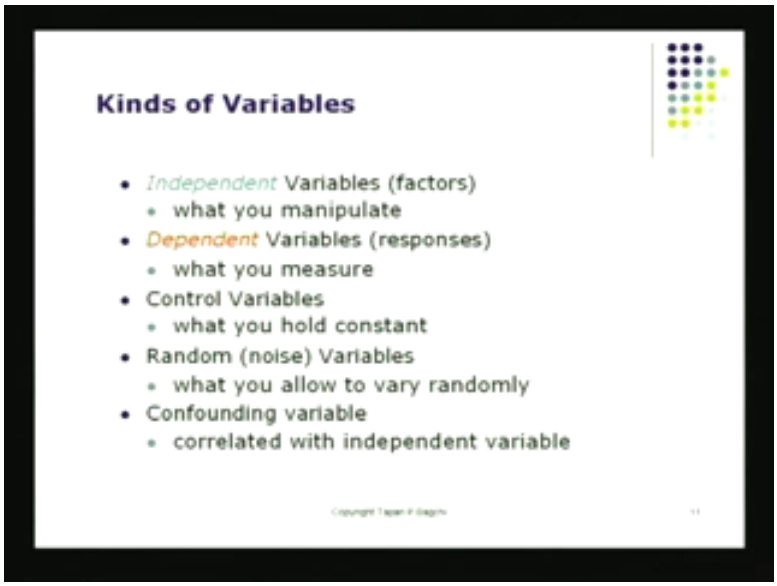
occur together, find to be occurring together for example, number of packs of cigarette smoke versus the extent of cancer contraction, the these are like these seen to correlated.

But, of course, something I should caution you about just simply observing a correlation does not really mean that one is causing the other, just simply because, I have got some data and plotted the data and I see this curvy kind of thing going up and I see then kind of going up together. This by itself is not the proof, that the cause is affecting the effect and a cause has anything do with the output correlation, the correlation relationship is not really guarantee that one is causing.

Therefore, what you have to do is when you manipulate the input, we then have to observe output and we have to really see then, after manipulate the input variable does that produce a corresponding change in the output, this is an experiment; this is not passive observation this is not passive observation, here we are deliberately manipulating a factor an experimental factor.

And then corresponding to each of those setting observing the output, then if I see a dependency, if I see a dependency then then I will say perhaps this is causing that, and this again I have to see in the background or noise that is there, because noise is also affecting the output therefore, what we have say is, it is just like my voice if my voice was so low, and the noise was so high in the room that you could not hear my voice, perhaps and what you to do is you have make sure this signal is stronger than the noise, only then you say that the signal really is significant otherwise it is not.

(Refer Slide Time: 34:29)

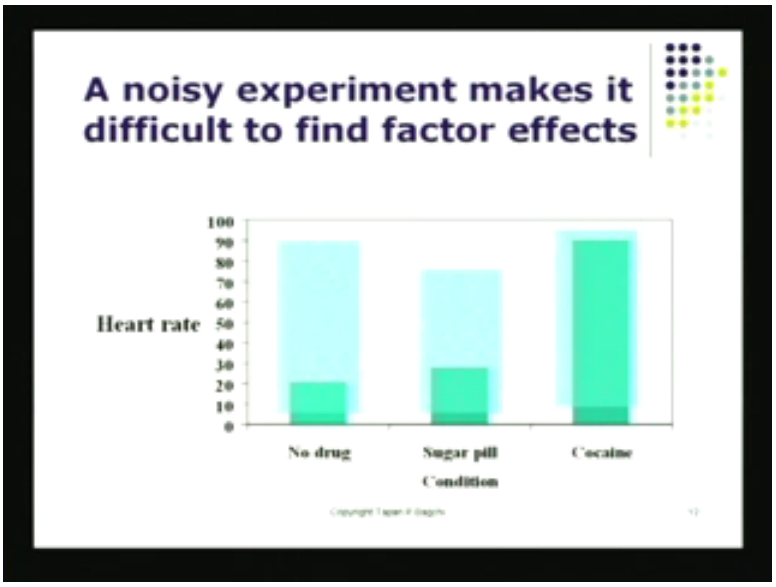


So, correlation by itself is not a not a sign for causality for that you have to conduct some special experiments, what kind of factors are we talking about of course, there is something called the independent variables these are factors that we manipulate in an experiment, and the dependent variable is the response, control variables are those those things that you might like to hold constant when a process is running.

And random variables are those factors that you are not controlling and no one is manipulating them and they are changing on their own for example, vibration and voltage fluctuation and so on; those might be changing on their own for example, and of course, something that is called a confounding variable is I have got one factor already set at a certain level in a particular experiment.

And I am putting another factor, in fact coinciding with the position of the first factor, we should have to we should try to avoid such settings and experiments, because then it will be very difficult for us to say, whether the effect was due to my left hand or was it due to the right hand, this is something we should try to, try not to get into, if we do that our experiment design will not be a satisfactory design.

(Refer Slide Time: 35:33)



What is a noisy experiment, on the x axis I have got different experimental conditions a person is given no drugs or he is given just sugar pills or he is given cocaine in an experimental setup, on the y axis I have got his heart rate I will measure his heart heart rate and notice here there are these dark green box and there are these dark green blocks and each of those settings there but, also there are these light green box what I really observe is not only the effect of the this dark box, but I observe this one.

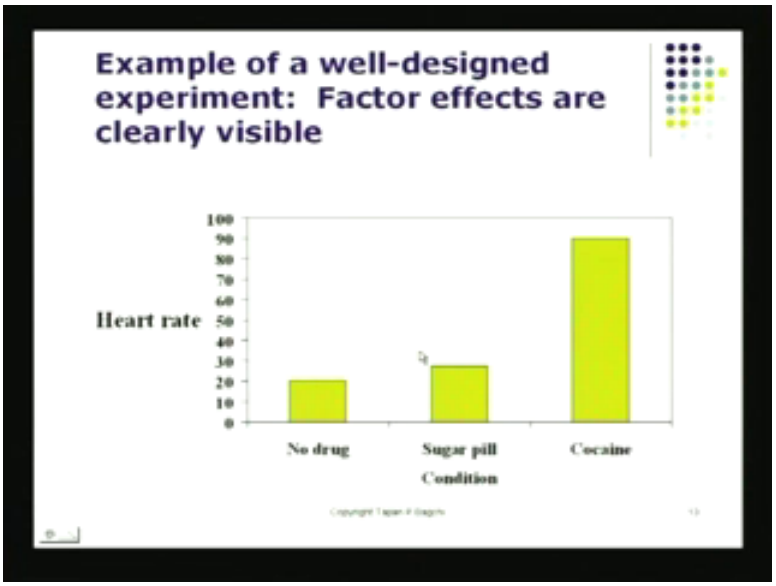
So, in fact, when I am watching from this side, I will say 90 for this response, I will say may be 80 for this response, and probably 95 for that response there, now what is the problem here, my experimental factors are these drug, no drug for example, or sugar pill or cocaine, these are my experimental factor setting unfortunately, while these experiments are going on the background is highly noisy.

For example, if it is a matter of heart rate perhaps, the person when he takes these medicines these conditions, he just see these things perhaps at the same time, he is also doing a treadmill run for example, or he is jogging outside is something like that, I have given him this pill and he has gone out he started jogging.

Now, the effect on his heart rate is going to be totally due to this pill that I have given

him, if at all there is an effect of it, but also due to jogging or due to his kind of running on a treadmill for example, such a background is a noisy background, if I am trying to see an effect and that say effect is produced possibly by a noise factor, I am not I have not designed my experimental conditions very well.

(Refer Slide Time: 37:27)



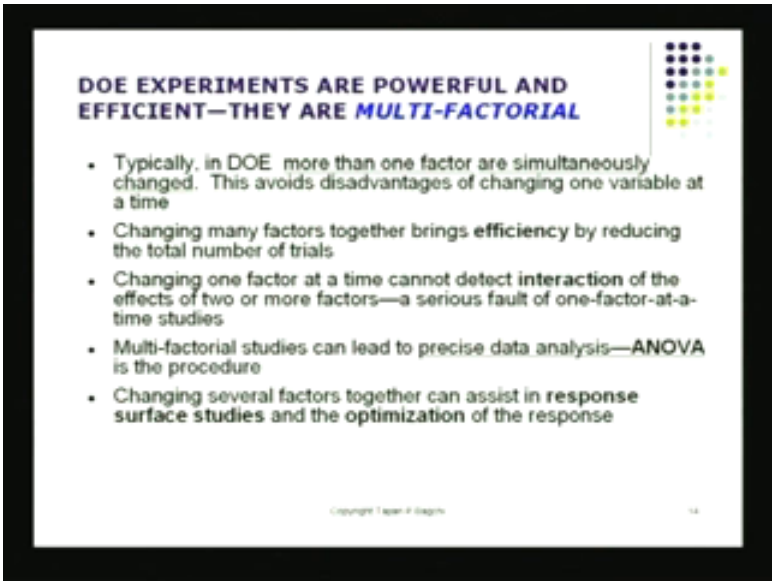
What is the choice, we will just control noise look at this condition, look at this experimental, setup this is a well designed setup, what what have we got there I have got here no drug situation sugar pill and cocaine, but I have controlled background noise very well, the result is that I am able to see indeed cocaine causes the maximum rise in heart rate, sugar pill does not do that much may be it does just a slight bit and no drugs of course, has hardly anything, it is just the resting time, whatever it is.

So, this is now, this is not a noisy background and I am able to see that, the lesson from this is when you are trying to conduct the experiments, do not run the experiments under these conditions, try to raise your experimental conditions in such a way, that you end up with results that are like this; except in a case and except and in a very special case when you want noise also part of your experimental study.

And I am going to give you an example of that, when I tell you about the design of

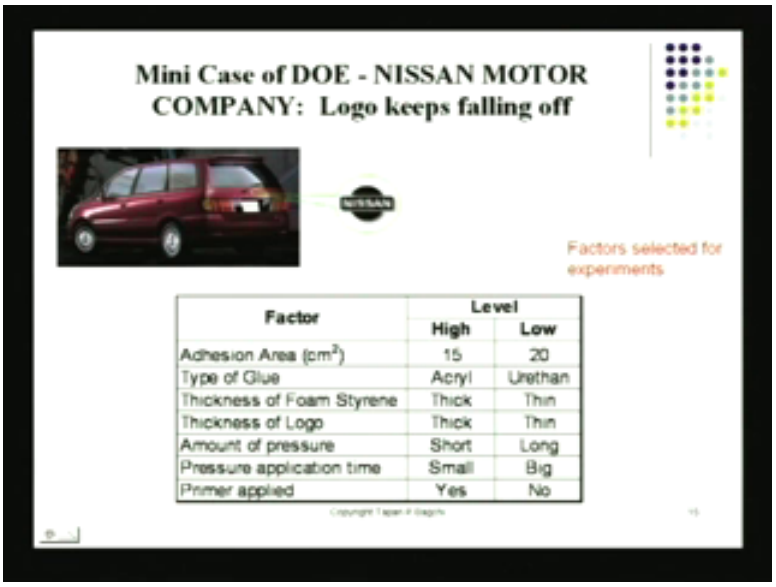
chocolate bars, chocolate bars are pretty fancy things to design and when I design the you know when I when I talk about the design of chocolate bars, you will actually see sometimes you would like to have noise also included, when I am doing a design study, when I am doing for example, some experiments I may include noise, that we will come back to in a little while.

(Refer Slide Time: 38:51)



So, in fact did the what we have what we have figured out, so far is we have got multi factorial situations that we must manipulate and typically we will be manipulating more than one variable, that is what we will be doing, we may change more than one factor, we may have their distinct settings there for data analysis we will be using an anova that is something that we would be doing; our ultimate goal is going to be to try to construct some sort of a response surface, so that we can come back and find the optimum settings for the experiment, without doing that we of course, cannot really say that we have really help the process.

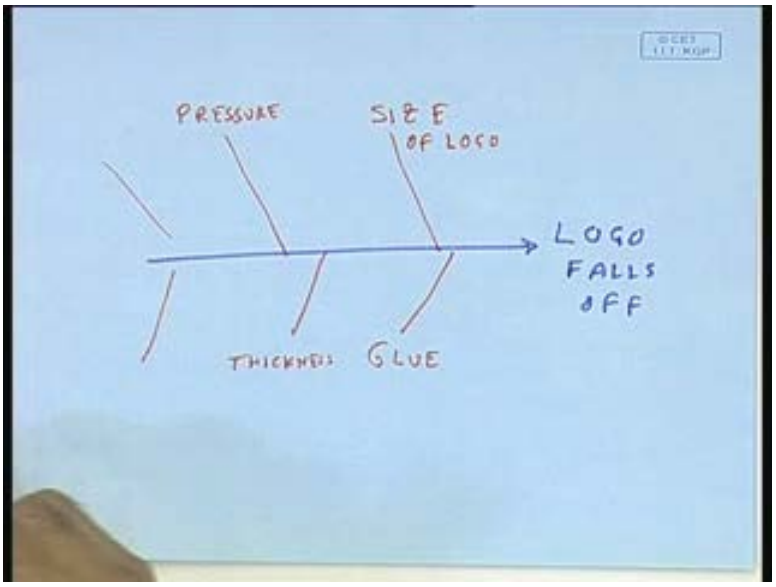
(Refer Slide Time: 39:21)



Let us take an example, let us take a very simple example, this is an example that is a six sigma case and the problem was this that many customers who bought Nissan cars, they found after a while that the logo fell off and you notice their logo there, the logo is right there, that logo is stuck with some pressure and some glue the logo is stuck on the back of the car. This little styrene there, it is stuck right there and this too go would keep falling off and this was a kind of a somewhat have a mysterious situation, people did not quite know, what was going on and they did want to kind of you know do something about it.

Now, one way would be to be just basically try different glues or different weights of logos or different sizes of logos, and it basically manipulate only one factor at a time, you could do that, but would that really work suppose I change the logo and I also change glue. What would happen, for this I needed to manipulate 2 or 3 or 4, 5 factors together and that is shown in the matrix which is here, this is the matrix that actually gives you an idea of the different factors that might be affecting that logo there.

(Refer Slide Time: 40:43)



So, let us just for a minute let us draw a little cause and effect diagram, I have on this plate here, I have got logo falling off, logo logo falls off that is the effect, and what are the different factors that might be causing this, I am again going to speculate I do not for sure it could be glue, it could be the size of the logo, it could be the thickness of the logo, it could be the pressure that I apply, any of these things are there may be many other factors also those might be effecting this thing.

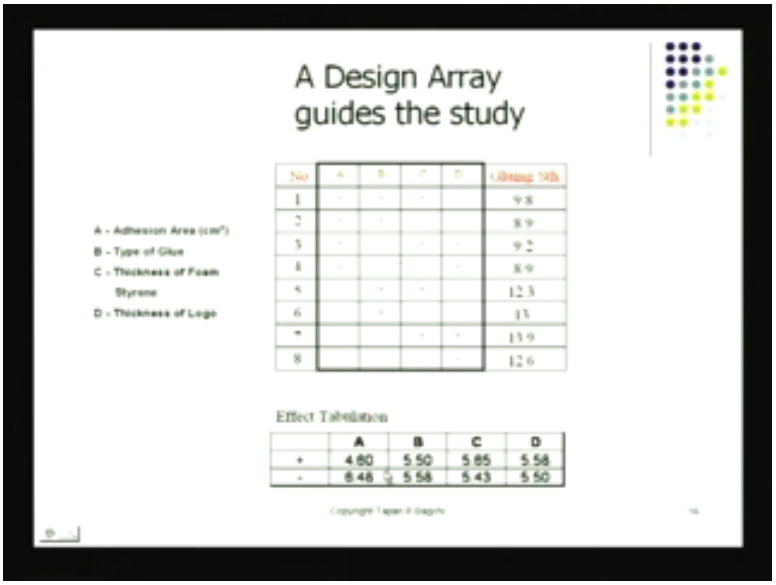
If I look at this situation since any of them, any of these and perhaps their interaction may be causing this problem there, what we have to really do is we have to conduct an experiment and for that we come back to the slide again and notice here these variables adhesion area, the area over which you apply the glue, type of glue that I use, thickness of the foam styrene which is their between the car and the logo, because the logo is a metal you do not want to damage the car.

Thickness of the logo itself, amount of pressure that I apply in trying to make sure, I stick the logo firmly enough and I let it dry, let it set then I remove my hand there. The amount of pressure that I apply pressure application time, that I just mentioned and any on a primer that you apply on the surface of the shining car which is there, that shine itself that that polymer that is there that itself, might resist sort of proper gluing.

So, that also is something that I might like to manipulate, now I would like to run a simple experiment and what is the kind of experiment that I am going to run, I am going to pretend that there are two areas available, 15 centimeter square and 20 centimeter square type of glue acrylic glue and urethane glue, these two types of glues are available to me thickness of the foam it could be thick or thin, the thickness of the logo it also could be thick or thin, amount of pressure could be short or long small and big so on and so forth.

Notice each of these factors have now, what we call treatments, two treatments for each of these factors, so if you now look at the look at the experiments that were run, what experts suggested that instead of trying all seven factors, we should focus perhaps on the first four and this seem to be the ones, that could be the culprit that could really fix the problem. So, for that, the first four factors were chosen adhesion, glue type thickness of the foam and thickness of the logo, having done this.

(Refer Slide Time: 43:32)



These setting too were taken out and they were put into matrix and notice the matrix there, this is the matrix this is the matrix, this matrix has these factors there factor A, factor B, factor C, and factor D and those are written up here, A represents adhesion area, B represents glue, C represents the thickness of the styrene foam, styrene foam that is

there and D represents the thickness of the logo and that is what I have got allotted there, I have got column A, I got here a there A sign, then B A sign on the third column and then I have got C a sign and D a sign.

Now, the first trial that I will be running, I will be probably be mounting may be ten of these logos under conditions which is like A at plus settings which is the high setting, B at plus setting that is a high setting, for a C at plus setting and D at minus setting.

And I run this trial and I measure the gluing strength, then measure this strength with which I can now remove that that that that strength, that is there I have to apply some force to take it off, I do the curing and then I try to take this thing off, and I measure the strength that is required force that is required in trying to take the logo off; that would have something to do with the with the lasting quality of that logo sticking on the back of the car; I do the same thing there, so gluing strength is actually the response and these are the manipulated variables.

In my second trial I do A at high level, B at high level, C at low level, D at low level and so on, and so forth. So, I run eight different trails I run eight different trails and observe the growing strength for, each of these things I have got for each trail, I have measured my glue strength notice here a couple of things this is a planned experiment, I have four factors that I am playing with, I have two settings for each factor.

So, I have got 2 to the power 4 in fact, I have got now 2 times to which 4, times 2 which 8, now in fact I have got eight trials here, but I am manipulating four factors and I have done a very special construct of the matrix, this matrix has been specially constructed, I will show you later, how you pick the right kind of matrix to able to do it.

Once I have done that, I can run my trials and I made my observation and this is like the the green fat arrow that was there, that was bring in data, so I have got my data now which is on the, in the last column; my data is right here, my data is in the last column there.

I take this data and I do some calculation, those calculations basically we show me the

effect of changing A from low sitting to high sitting, and look at the impact glue changes, I do the same thing for B, I do the same thing for C, I do the same thing for D, these calculations had been done in a special way, by starting with these input conditions and by looking at the construction of the matrix, the structure of the matrix; once I have done that I am able to calculate, I am able to generally I am able to calculate these factor effects there. Now, what this is showing is, when I change from plus to minus these are like two settings for a glue instant changes, it has the same impact on glue instant.

(Refer Slide Time: 46:46)

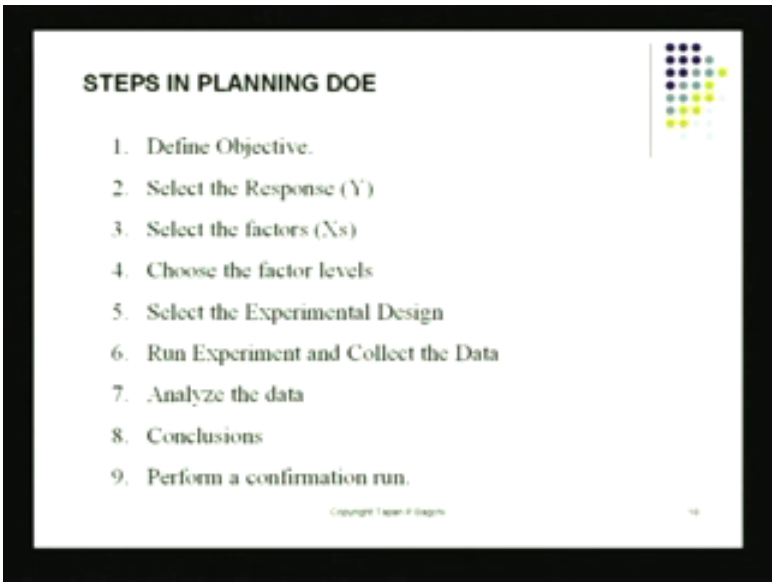
It may plot it up the plotting goes like this A is going from you know high to low in this particular case and the effect is that my gluing strength goes up, when I change B this is the effect, when I change C this is the effect, when I change D this is the effect. Now, by doing those eight trial I conducted exactly eight trials, and eight is like this, with those eight trials I have been able to study four factors and each of those four factors, they were set at two levels and my data manipulation technique by data calculation technique.

Let to these plots let to these plots, that you see there what does this tell us while as far as this gluing process is concerned, additional area is the is the one that is really the culprit, perhaps we were not using sufficient additional area there other factors seem to have an effect that is more less flat; so I need not really worry much about manipulating those of

the factors there.

If you really want to we could actually take them to the level where they give us, they they would give us a high strength that we could do, but as far as fixing the process is concerned, if you would change the setting of A from minus to plus immediately give us a big jump in gluing strength. This came out as the result of doing the experiment and the red circles they show us the optimum settings now, for basically running the process basically running the process which is there.

(Refer Slide Time: 48:09)

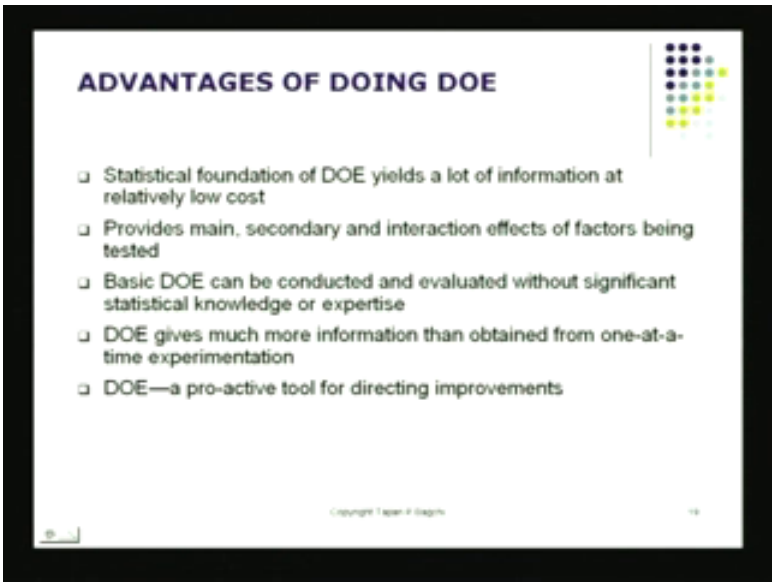


What are the steps in conducting the experiment just take a look at, how we conducted the experiment, we stated your objective then we also selected the response, then we selected the factors of you wanted to manipulate, we choose the factor levels for each of these factors the x factors, we selected the experimental design which is the matrix I have been talking about, then we ran the experiments in a randomized order and we collected the data, then we analyze the data and we get got those plots done and we do some conclusion.

And of course, once I come up with a with an empirical conclusion, I still have to conduct what we call the confirmation run, I have to convert one conduct, one

experiment under optimum condition. So, for example, for the gluing experiment I should run one experiment that sets A at minus, B at minus C at plus and D at plus, I must run one trail to basically confirm what we have deducted what we have deducted basically, what reduce empirically, that is something that got to be able to do.

(Refer Slide Time: 49:26)



What are the advantages of doing it this way a lot of advantages, I could produce a lot of information by the conducting experiment, and this is like something that I was able to do without too much trouble. I could obtain the main main effects, I could also work out the secondary effect if there is any secondary effect, I could also work out at also calculate the introduction effect; now this could be done for each of the factors that are been manipulated, and in fact it turns out, because of matrix are very simple to use and the settings are all clearly laid out in the matrix.

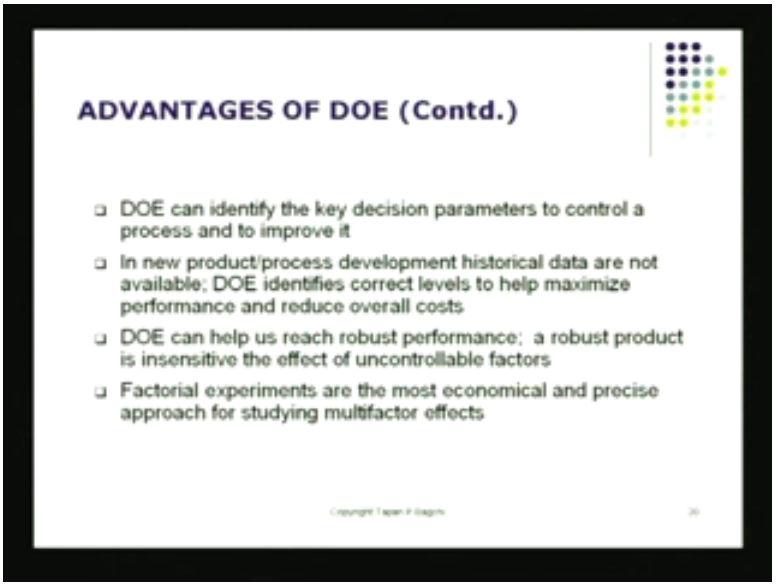
So, anyone who really understands how to manipulate the variables, you can take one of those matrices and run the trial, so running the trial does not really require you to know lot of statistics that is something very easy in modern experiments. In the older experiments, you also had to have to do something called anova, and its variance to come up with with your optimum settings and so on, so forth. Now, you do not have to do that, you do not have to really worry too much about significance, if you control the

background noise, and if you use a simple matrix you should be able to work out the factor effects and their interaction that you should be able to do.

And this is far more useful than one factor at a time experiment, which we are going to see again as we go later on, later on as we go deeper into this topic, you will be able to see that DOE produces information that is substantially better then the inversion that are produced one factor at a time, for one thing something that the one factor time experiment will not be able to able to tell you is whether there are effects between, whether there are interaction between two factors that can be manipulating independently. That effects point interacting some complex way, and this would not come out if I would come out one factor a time experiment that is like something that we got be able to keep in mind.

And DOE is also pro active too this is something that I could in R and D, I need not really go out and wait for production to began and the real plant is running there and I start my experiments, my my development work there, that is actually not necessary, that is actually not necessary, I could basically run all these experiments right at the R and D.

(Refer Slide Time: 51:48)



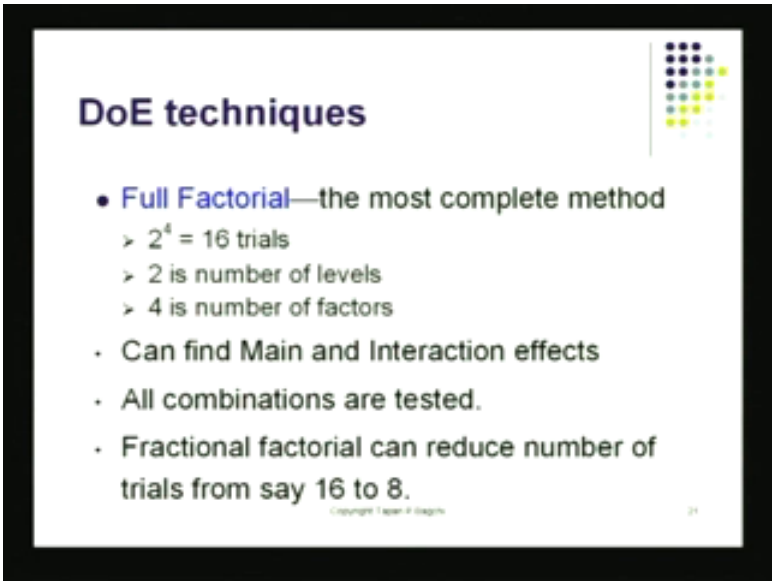
And really they would not be much trouble at all, so other advantages are also there for

doing DOE, it can identify the critical precession parameters to control a process and to improve it. In fact, we saw that factor A was the most significant one the others did not really matter, as far as fixing that the dots and problem was concerned, that we could see quite easily.

New product development when you manipulating many different variables DOE is a is a short cut way, DOE is an efficient way and it identifies the correct levels to help maximize performance and also to reduce defects and the overall cost, that also can be done if I set up those as a objectives of the experiment, I am able to those quite easily.

DOE does something even more DOE does robust design, DOE helps us to robust design and this is the method that was used by Taguchi and that connection again I am going bring back the chocolate bars, I am going to tell you exactly how to construct robust design, that is robust chocolate. And that property I am going to keep a little secret right now, till I start discussing Taguchi methods that I will do, and the factorial experiments they are turn out to be most economical, and they are also very precise.

(Refer Slide Time: 53:01)



So, there are lot of advantages there and there are of course, there many different techniques and one of the popular techniques for doing experiments is to use the full

factorial design, this is a most complete technique, it can find a main effects and also it can find interactions, so both of those things can be found if you use full factorial designs.

All combinations of the different treatment levels those can those can be powered by doing the full factorial trials that is there but, if these experiments become too many I have to use some other method which is called running a fractional factorial run, and that I will that I will show you once we get into that, I will show how we come back from full factorial design to partial factorial design.

That is actually pretty efficient way to find a lot of information without having to run all the different to the 2 2 to the power k experiments, which is like the full factorial design, we will continue with our lecture in just a little a while, thank you very much.