Strategic Performance Management
Prof. K. B. L. Srivastava
Department of Humanities and Social Sciences
Indian Institute of Technology, Kharagpur

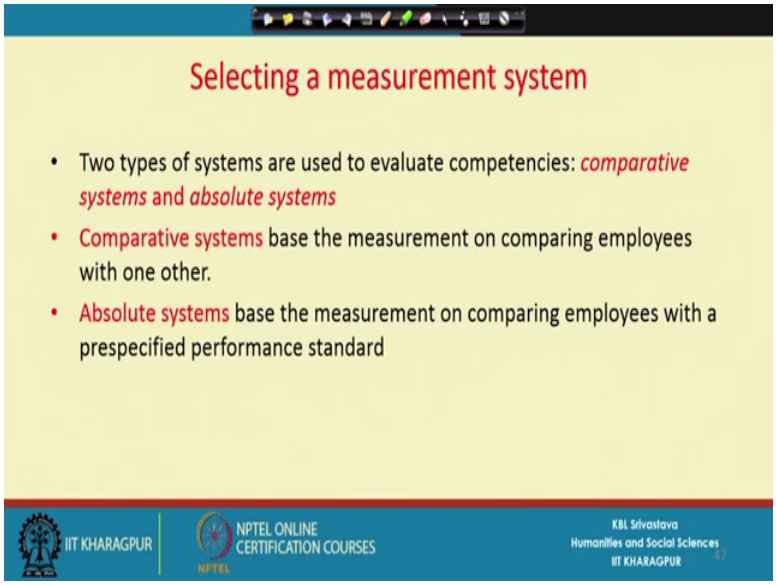
Lecture – 14
Measurement system

Welcome back to in the next lecture that is 14, and in this lecture we will talk about that how to measure performance in terms of behavior and standards. So, you remember we talked about performance, standards. And three things that we talked about is: identifying key accountabilities, a specifying objectives, and then looking into performance standards.

Now, once we have done this job the next job is to see that how we are going to measure behavior and performance. And there are two kinds of measures that is available: one is subjective which is used for measuring behavior, and the other one is objective which is used for measuring results. But here our focus is more on measuring behaviors, because for results they are predefined specified targets and standards against which the performance is measured.

So, our focus here is on measuring more on behaviors and less on results.

(Refer Slide Time: 01:16)



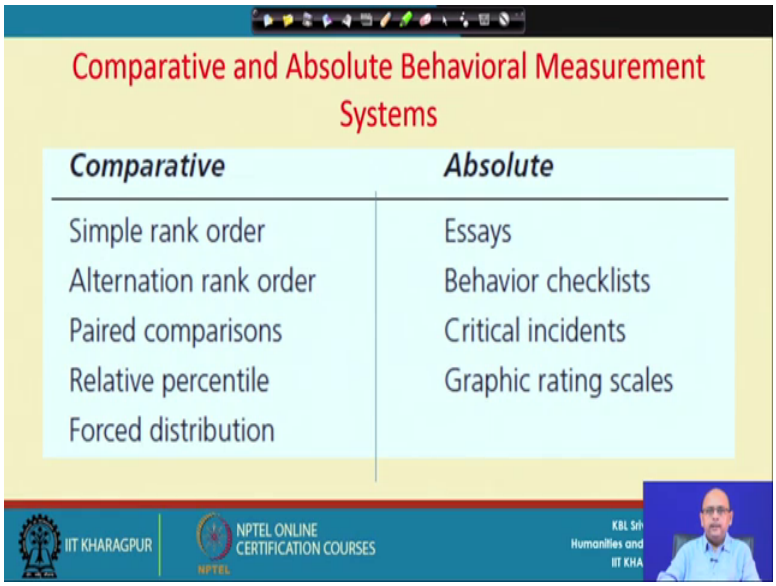
Now, when it comes to identifying a measurement system that what kind of measurement approach you are going to use there are two kinds of systems that is available; that is comparative systems and absolute systems. In a comparative system what happens you try to compare one employee with another employees and see that whether that employee has then better or worse than the other employees.

And here, in absolute system what happens that you try to compare using certain pre-specified performance standards. For example, when we are going to evaluate students we can use both kinds of systems: comparative systems and absolute systems. Comparative systems are relative in the sense that we look at the performance of one student compared to other students. And that is how you evaluate them, right and when the grading is based on this relative system.

Now when we are talking about absolute systems we already have pre-specified standards against which the performance of each a student is compared. For example, in a comparative system suppose the highest score is 60 out of 100. So, the person who is getting highest would get a grade, right. And accordingly the grades of other individuals are decided depending upon the average performance of the class. But in absolute system what happens we do not compare the performance with each other, but what we do we have already specified performance standards. Like we say that if you (Refer Time: 02:46) is the mark again above which you get a grade, right. 70 is the mark against which you will be compared and if you gets 70 and above you get B grade.

So, what happens in this case? You are following an absolute system where you are comparing a individual against a pre-specified performance standard and you are measuring his performance, right. So, this is the difference between comparative systems and absolute systems.

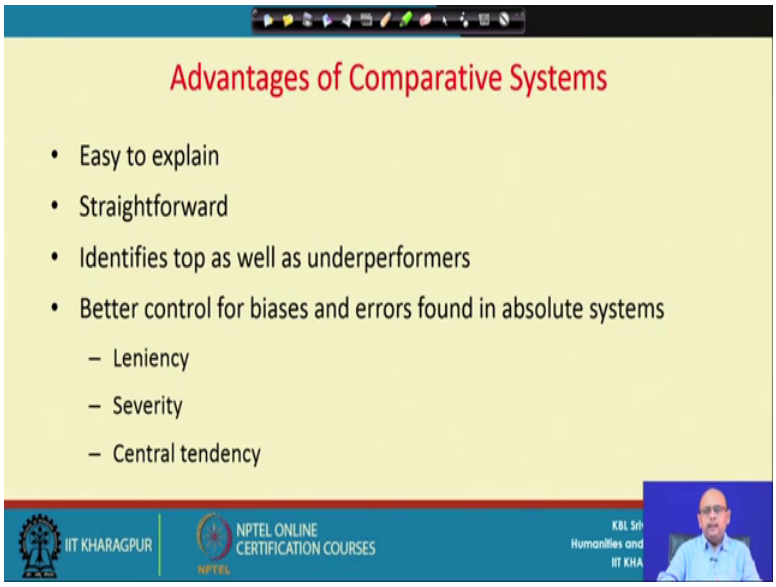
(Refer Slide Time: 03:15)



Now these are the methods that we use to measure performance. In comparative systems we have rank orders, alternative rank, paired comparison, relative percentile, and forced distribution. But when it comes to absolute measures we have essays, behavior checklists, critical incidents, and graphic rating scales, right.

So, what we are going to do is that discuss each and every measurement system whether it is relate to comparative or absolute system one by one.

(Refer Slide Time: 03:44)



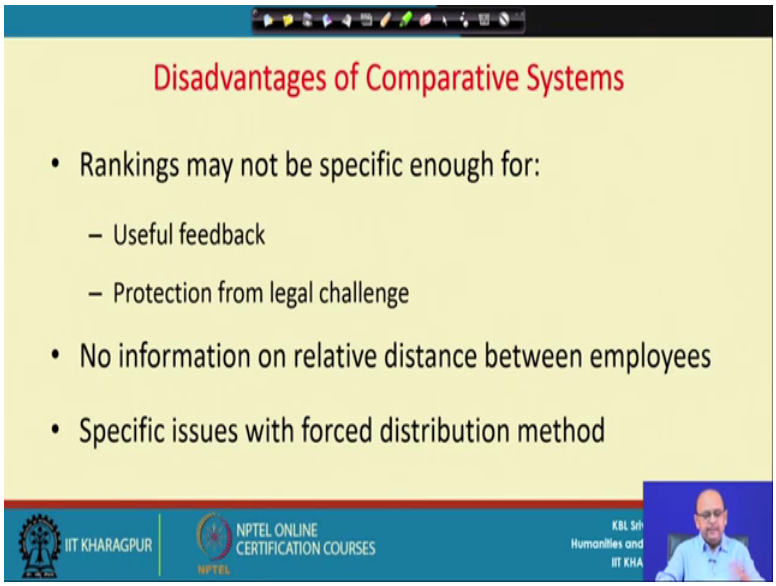
Now, one of the advantages is that of comparative system that you can better understand, because you are going to compare one person with other person. And here you are neither biased nor you are making, because you look at the average performance of the class.

Now, here you are comparing the performance of one person with another person. And accordingly you decide whether the performance is good or bad; so that is; what is the advantage of comparative systems. But in case of absolute systems which is not there, because once you have identified a standard you have to go by it irrespective of what is the characteristics of the sample of the class.

For example, suppose everybody is very good in the class, right. So, you are going to use an absolute system what happens everybody will get a A, B and above. But in a comparative system what happens you will be able to differentiate the performance better. So, you will able to differentiate good performers from bad performers, right which is not possible in a absolute system because everybody above 80 gets A grade, right. So, you are able to control certain things relate to leniency, severity, central tendency, right.

Because these are the things which are going to affect your performance, which will discussed later on. Leniency is become you are not, you are going to be too lenient are your going to be very serious very strict or you are giving an average rate to everybody, right that is avoided.

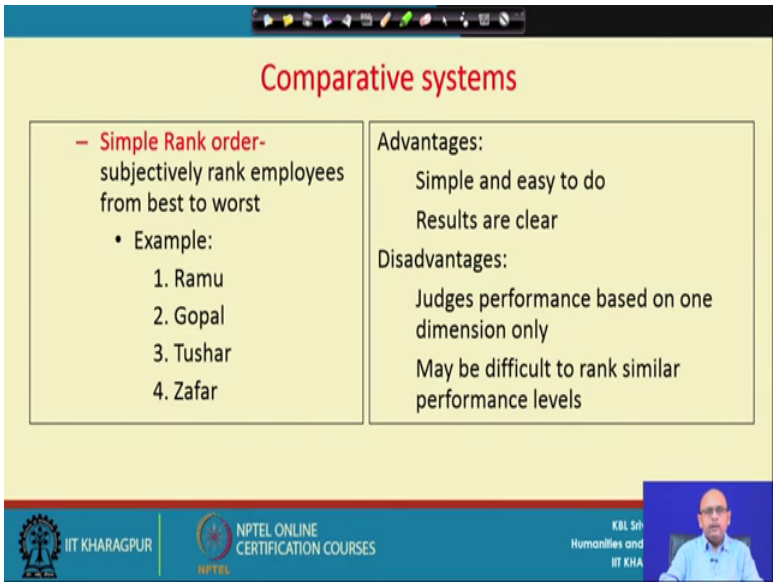
(Refer Slide Time: 05:22)



Now, the disadvantage is basically that you can be challenged, because you are not measuring it against a standard, right. Another thing is that you do not know how relative the score is from one individual to another individual. For example, suppose somebody has got 60, but somebody is got 57. Now between 58 sorry, 58 and 60 what going for a little grading you are giving A to 60 and B to 58 this makes huge difference, right. But if you are going a absolute grading both of them will get the same grades, right.

So what actually happens, if you are going for forced distribution then this is a problem right. So, you do not know; what is the distance that you need to maintain in order to differentiate in the performance of the paper. So, this could be one of the disadvantages.

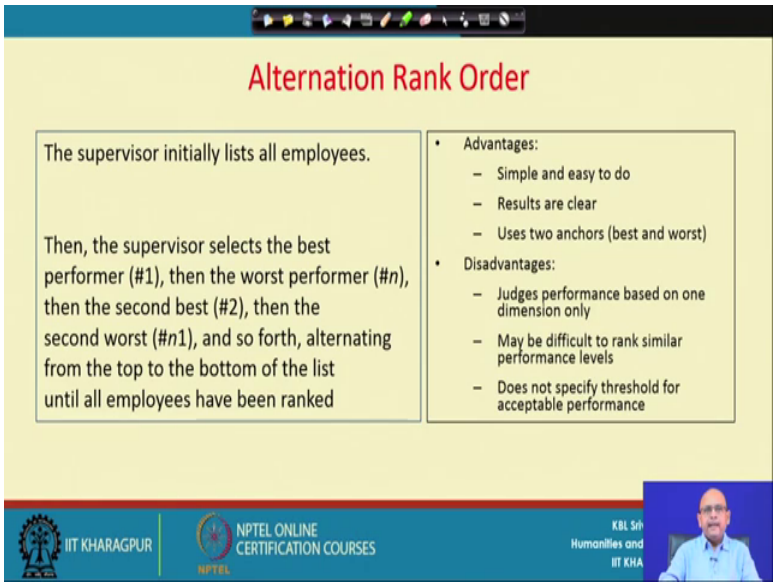
(Refer Slide Time: 06:23)



Now, as I told you I am giving some examples here to explain different kind of comparative systems; in what happens this in simple rank order. What you are going to do here is you are going to subjectively rank employees from best to worst. So, how it happens? It is based upon the subject matter experts and their judgment which is very very subjective in nature. So, there are five people say- Ramu Gopal Tushar and Zafar. So, what you are going to do here? You are going to rank them based on your judgment which one is the best, then the next best, then the next one, and then finally the last person. So, when we want to compare the performance I using a simple rank order you say that Ramu is the best performer and Zafar is the least performers, right. It is based on the ranking depending upon what their performance is which is just subjectively, right.

So, what are the advantage of this? It is very simple, results are clear you can very see that who, what is the rank of the employees and accordingly you can identify who is the best performer right and who is the next best, but there are certain disadvantages. This judgment is based on just to an dimension. There could be other things which are not considered here, but it is not possible to consider in there. That is why simple rank order is not used much, right. And what we are doing here only was rank in the performance of then there; this is a results you are not considering the behavior, right. So, these are the problems of the simple ranking.

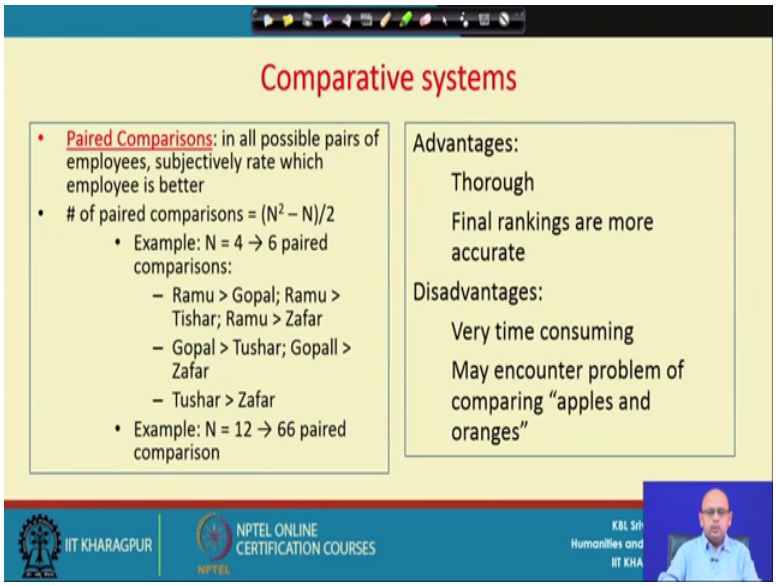
(Refer Slide Time: 07:57)



Now, in addition to that you adopt a different method that is not a alternative ranking, right. Alternative ranking what happens; so you will have a list of all employees then supervisor the select the best performer, then and again the worst performers. So, first of all what you do you; identify who is the best performer who is the worst performer, then the second best, then the third best second worst. So, who is going to be number 2, who is going to be number 4? So, alternating from the top to bottom the list of all employees is prepared on the ranked.

Again it is simple and results are clear. And you are having two anchors best and worst, right. But again it is based on the dimension only. So, those who are having similar performance sometimes it is very difficult to rank them. And you are not able to specify the threshold for acceptable performance, because that could be use for taking certain administrative is development and this universe.

(Refer Slide Time: 08:57)



Third one, that is the comparison; paired comparison.

So, what happens in paired comparison? Suppose there are five people. So, you are comparing everyone with each other, right. So, if you look at the previous example Ramu, Gopal, Tushar, Zafar everybody is; right Gopal Tushar Zafar Ramu all of them are going to be compared with each other. So, you are going for paired comparison. So, the number of paired comparison is going to be N square divided by N by 2.

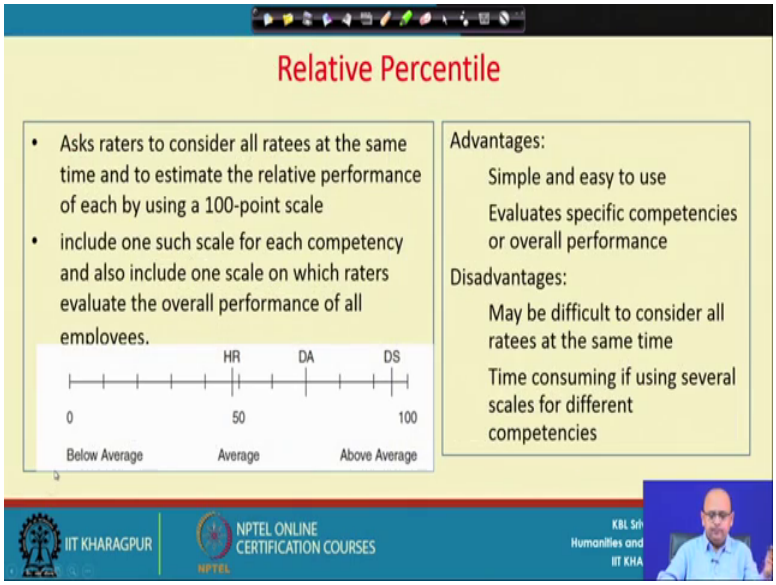
So, if there are 4 people you are going to have 6 pairs, right. So, the comparison is like that: Ramu is compared with Gopal, Ramu is compared with Tushar, Ramu is compared with Zafar. Now Gopal is compared with Tushar, Gopal is also compared with Zafar, and then finally Tushar is compared with Zafar. So, if you are having 4 people you are making 6 comparison, and then you again go for ranking only because it is a ranking but what actually happens here that you are going for paired comparison; we are going to pair it with everybody, each one is paired with everybody.

Now, this creates a problem provided there is a more number if there are few people you can compare each other, but suppose there is 12 people. So, if you are having 12 people then? The total number of comparison goes up, like you have to you look at the formula that is N square minus N by 12. So, N square means 12 is the number, so 12 square 144 divided by N 12 how it comes it; it comes out 32; 132 divided by 2 so it comes out to be 66.

So, what I am trying to tell you that if you increase the comparisons then it is very difficult to do this exercise, right because it is not possible to compare say more than 10 people at the same time. Suppose the 100 people. So, if there is N is 100 then what happens? How many comparisons you are going to make? And the process you are lost you are; it is very difficult to rank them properly, right.

So, if you have having less number of people then it is good, but if you are having more number of people then you in compare certain problems. That is like looking at apple and oranges; who is good who is bad when to put them, right. Similarly it is very time consuming time; time consuming in the sense that if a so many people then how would you to carry out the sacrifice. So, this is another problem.

(Refer Slide Time: 11:35)

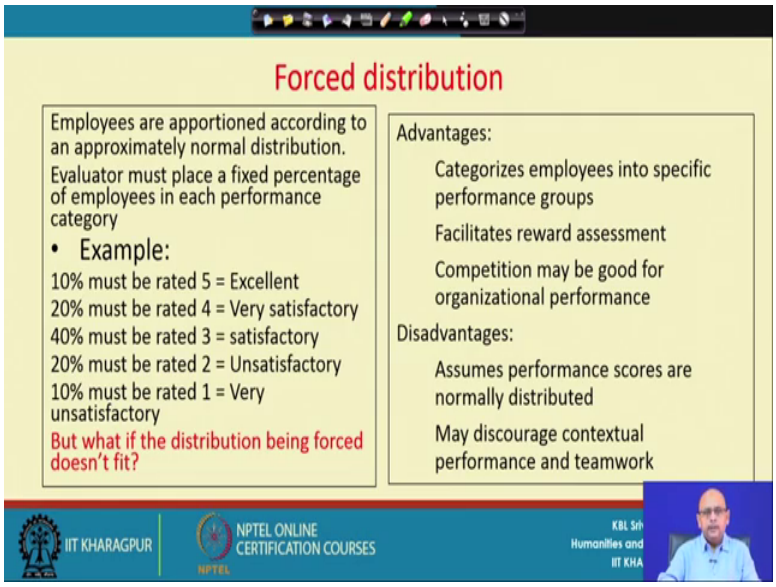


Then another method is relative percentile, right. Relative percentile is that you ask evaluators to consider all rate is at the same time, and to estimate the relative performance of each using a 100 point a scale.

So, there are three people right Hare Ram Dayavathi Dhamodar suyog; so this three people are there. Now if you look at this, you are going to use a scale from 0 to 100 and were going to put them into three categories. So, you have put HR as average, DA is between average and good; average and above average, and DS is what you call above average, right. So, for each competency have one scale. And you are going to rate are evaluate what you call the overall performance on that competency, right. It is good to

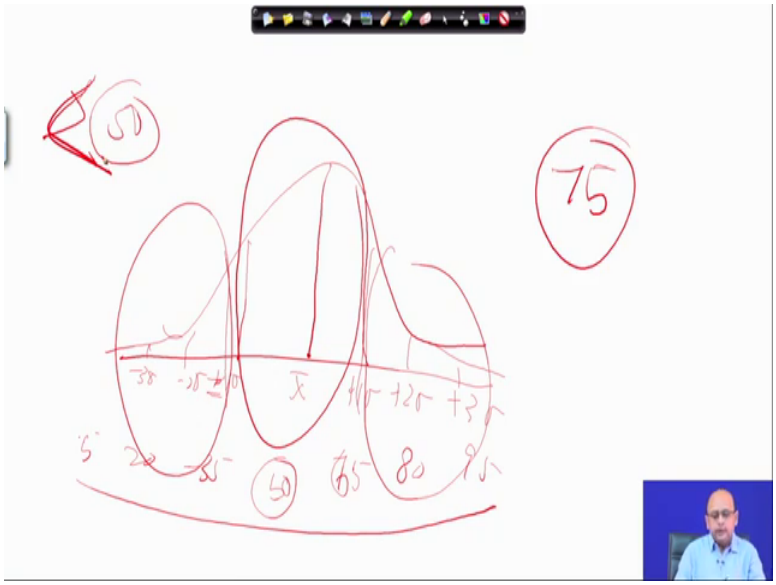
use and each competencies; for each competencies you are going to have one scale on which is going to be measure, right. But difficult to consider all rate is at the same time, because you are 3-4 people are there then you can rate them, but suppose are 100 people for which you are going to rate on each competency then it is a very cumbersome job and there time consuming also, right. And for each competency you are going to have different scales.

(Refer Slide Time: 12:57)



Then coming to another scale this is known as forced distribution. In case of forced distribution what happens? Employees are evaluated based on normal distribution. So, let me give an idea how it happens. When I am talking about normal distribution.

(Refer Slide Time: 13:18)



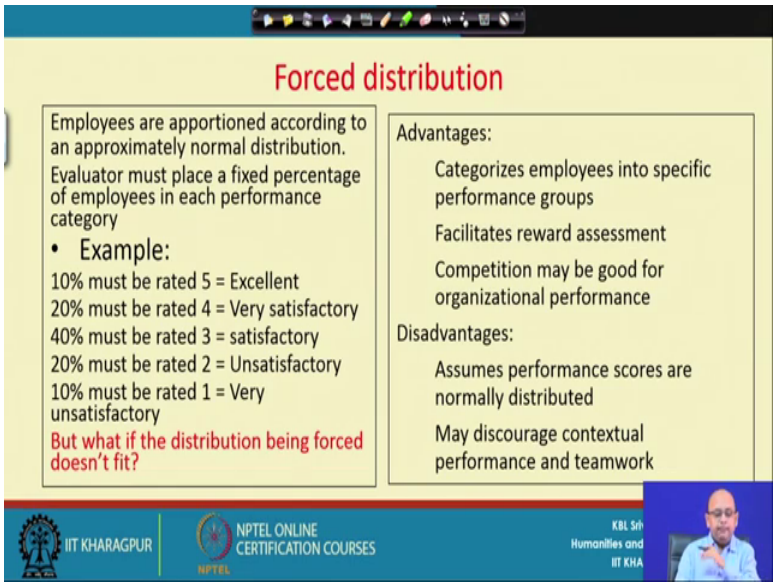
This is a normal distribution right which is plotted got it to 3 sigma plus 1 sigma plus 2 sigma plus 3 sigma, right. And this side also we have plus minus 1 sigma minus 2 sigma and minus 3 sigma; the two ends never made. So, this is the normal probability curve, right.

So what you do: this is you treat as average performers, this is what you treat as good performers and this is what you treat as bad performers. So, what actually happens? In the process you try to use this distribution and forced employees to be put into one of these categories which may not be good. Why, because if everybody is good then; and you are going to force them then how would you force them into different distribution systems. If everybody is bad then how would you identify them to be put into higher categories, right?

For example, suppose you say that 50 is the average score on which you are going to performance. So, say 65 80 and say 95; this side it is 35, 20 and say 5; so the range of the courses from 5 to 95. But suppose everybody gets 75 and above then. And you are using a forced distribution to put them into different categories, so this is a problem. Suppose everybody is getting less than 50; sorry, less than 50 then how would you force them into different categories?

So, this becomes a problem if you are going to use a forced distribution system, right.

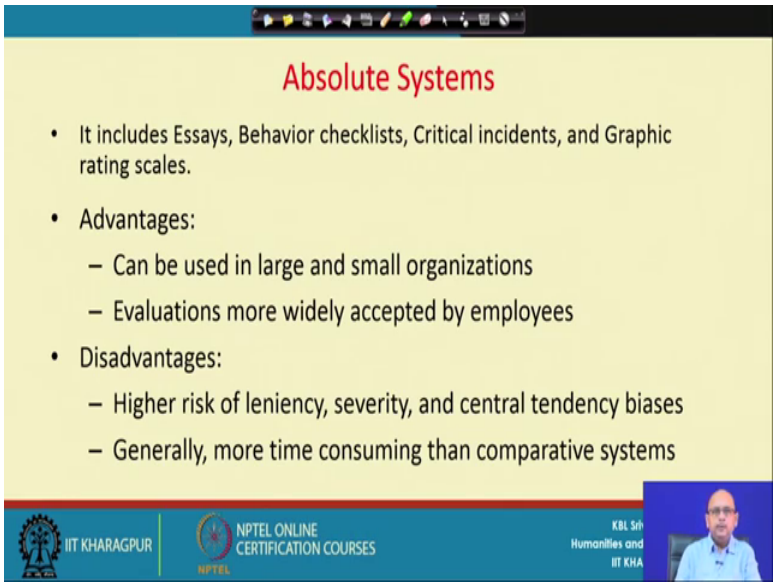
(Refer Slide Time: 15:57)



So, in a forced distribution what happens? You will need to rate some people as excellent, some people as bad. So, for example look at here 10 percent are rated excellent 20 percent treat as very good; 40 percent as good, 20 percent again not very good and 10 percent as bad. As I told you that everybody gets scores a 75 and 80 and above then how would you distribute and force people to be part of either excellent, satisfactory, or very good. As suppose everybody is bad then how would you force them to be in excellent? So, this is why forced distribution is not very very effective in the process, right. But many companies is they want to promote very few people what they do they use this distribution and try to promote only those who are in the top 10 percent bracket; treating them as are considering them excellent. So, the far not going to be promoted, ok.

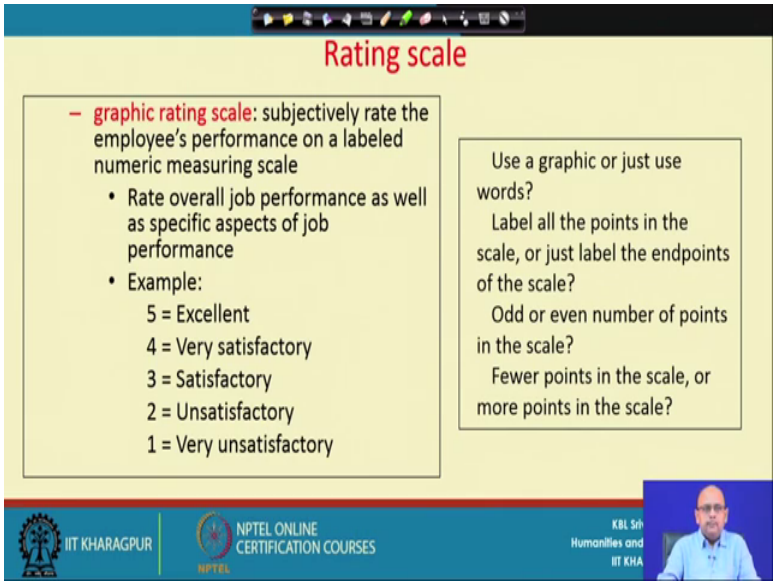
So, basically you are trying to categorize people into different performance groups. The idea is to reward them those who are doing very well, right. But they are go to be high competition? Suppose everybody is very performing very well then what will you do are everybody is performing very well then; it is possible to do only if the performance is scores are normally distributed, as I told you. But if the performance distribution is not normal then, it is not possible. And this discourages this contextual performance because you are looking only at task performance, right. So, it is not very good.

(Refer Slide Time: 16:52)



Then, from relative systems you are moving to absolute systems and in absolute systems what happens. We are talking about four different kinds of thinks that is: essay, behavior checklist, and critical incidents, and graphic scale. The advantage is that you have absolutely standards against which the performance is measured. And it is more accepted by employees, right. But the disadvantage is that you may be lenient, strict are you great everybody is average, right. And it is more time consuming then comparative systems. And when these absolute systems are developed they are based on certain parameters. So, that is very very important to understand.

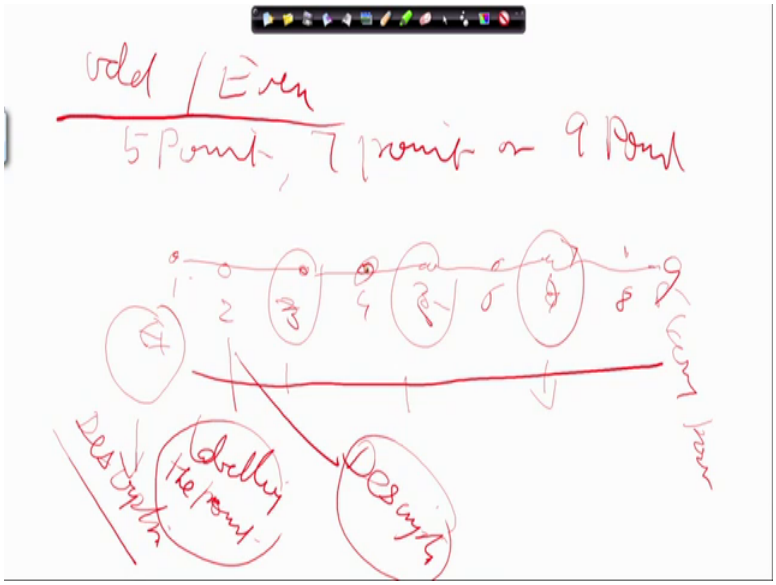
(Refer Slide Time: 17:34)



Now, look at it. Rating scales is a most common form of rating by the employees, sorry by the organizations among the employees, right. So, what you do when you are going to rate them basically you are looking at the performance and you have labeled numeric measuring scales, like performance is measured against these 5 point scale, right. So, you are measuring oral your performance or a specific aspect of the job performance using this example. 5 is very excellent; so perform performance is very good we get 5 rating; if we performance is very unsatisfactory they think gets 1, right.

So, this is how rating scales are used but there are certain problems, whether where using only graphic or just words. And whether going to label all the points are in the scale and label only the end points. Suppose you the rate only 1 and 5; this like this.

(Refer Slide Time: 18:35)



See for example: this is a scale 1 2 3 4 5, right. So, it is defined 1 2 3 4 5 right. So, will 1 is the excellent and this is very poor. Suppose you identify only these point and these points are not identified then. If there is no disruption then it is very difficult to do it, correct.

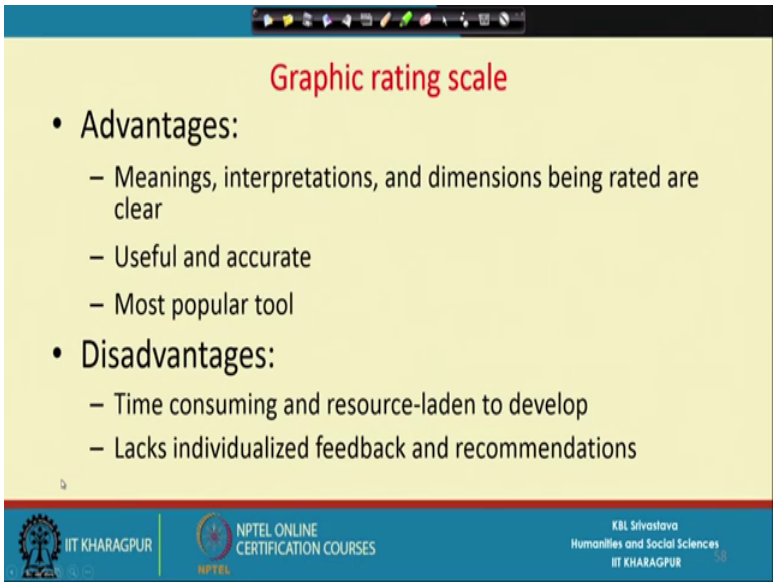
So, you need to ensure that you are going to provide label to each one of them. And then important issue is that what does it mean to say excellent performance. So, you also not only give this, but also provide a description; what does it mean to say by excellent performance; because, if you are saying excellent performance it means to say these things, because it would help the rater to evaluate the performance of for it in a better

way and the ratio. How many points in the scale should be there? Whether it should go for 5 point 7 point or otherwise say 9 point? As you increase the number of points, suppose you say 1 2 3 4 5 6 7 8 9 what are the problems? As you increase the number of points in the scale this creates more problem. And these problems are related to what? How do you label them? This becomes a problem.

So, labeling the point is a problem, providing a description to these labels is a problem; though it makes the job of a rater easy, but how would you label if you are increasing the number of points and the ratio is there. Which is, whether you are going to have an odd point to scale or and even point the scale, right. If you are going to have an odd point to scale you can have a neutral point, right. For example, if a 9 point scale would be there neutral point. Another issues related with it is going to be ascending order or not, or whether you want to have a neutral point.

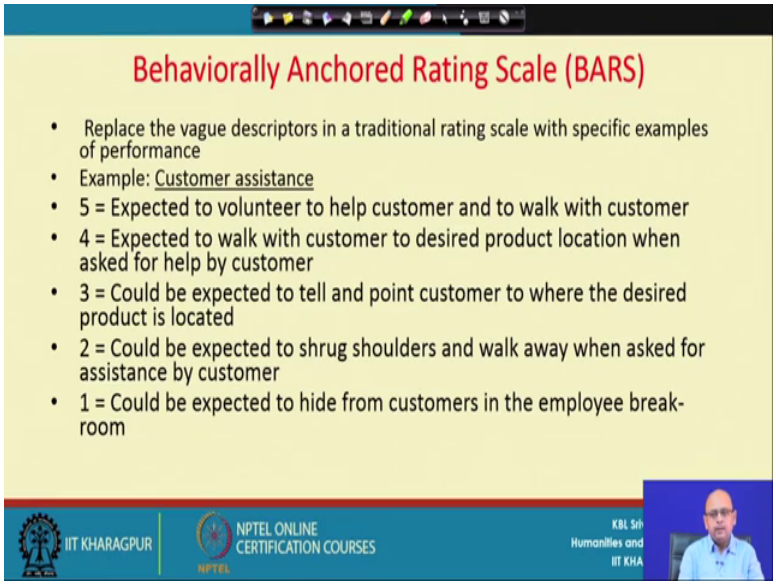
So, in most cases we try to avoid even points number because you do not have a neutral point in that case, so it is x. It is generally recommended to have an odd point a scale, not go beyond 7 point scale because it is very difficult to provide label and description to more than 7 point scales, right. So, this is what we wanted to discuss it. That is there are issues, but they were going to label all the points are only that top two points the first and the last point; whether you going to use odd and even number of points in the scale whether you want to have two points or more points in the scale. So, these are some of the issues writing a scale, right.

(Refer Slide Time: 21:34)



Now, another advantage of this is graphic rating scales, where you are going to provide label and interpretation of each dimensions are very very clear. It means graphic rating it is a advantage of rating scale that for each label that were providing you provide a interpretation and a meaning. So, it becomes more important and meaningful, right. But going that is very very time consuming and it will the requires lot of resources.

(Refer Slide Time: 22:01)



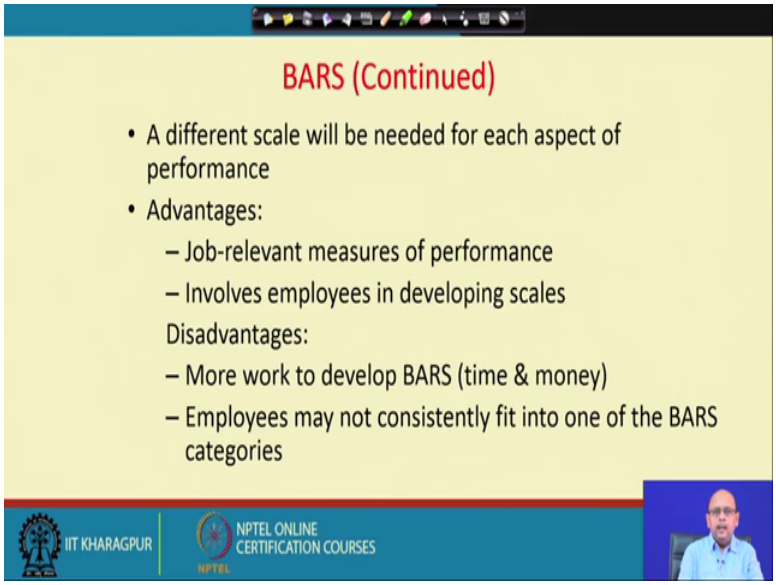
Then coming to another important scale which is a alternative or you can say, is a more refined rating scale this is known as behaviorally anchored rating scale. In case of

behaviorally anchored rating scale what happens? You are going to replace the vague description of a traditional rating scale with specific examples. Say for example: in case of a restaurant there is a waiter and he is supposed to provide assistance to the customer. So, this is the one competency that is to be measured on a 5 point scale. So, look at the 5 point scale.

So here what we have done, you have try to identify the behavior of the waiter on along these points and each behavior are each point you are very clearly describing the kind of behavior that is happening. Though it helps the supervisor the rate to better evaluate the employees, but the problem is that writing this kind of descriptions for each of these competencies is very difficult. For example, when you are saying customer assistance and you say that expected to volunteer to help customers he gets better rating, but you say that could be expected to structural as and walk away when asked for assistance by the customer; so this is the bad rating.

Now, another issue is that whether the waiter is making all these behaviors are not, any of these gives, and he is making a different behavior and still successful. Now the problem is that whether you are able to capture the entire range of behavior related to customer assistance or not, and then making a scale. So, this is a big problem.

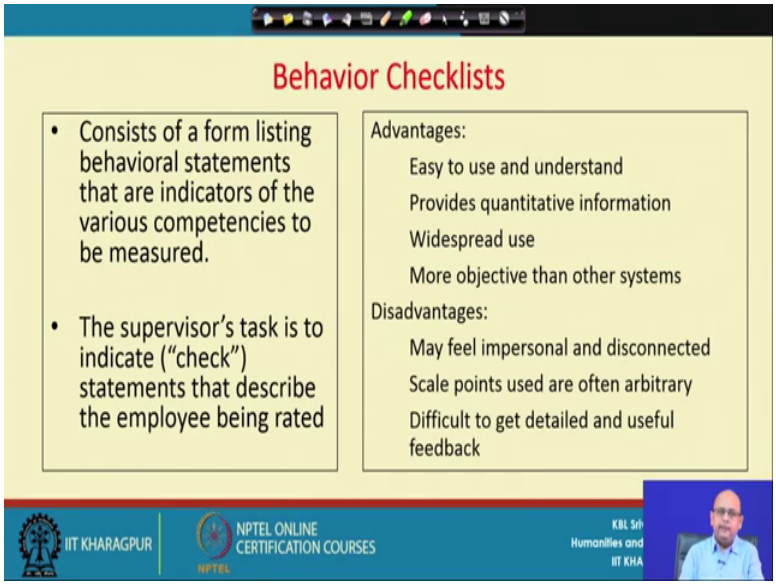
(Refer Slide Time: 23:42)



So, the idea is that is you need a different scale for each category. It is good if you are able to develop job relevant measures and you are able to develop a good scale then is the best thing to do.

But it requires lot of time and money to develop these kinds of a scale; and employees may not fit into the categories that you have identify. For example, he might be doing something else but is still very successful, but that behavior is not captured, right. So, these are some of the problems related BARS.

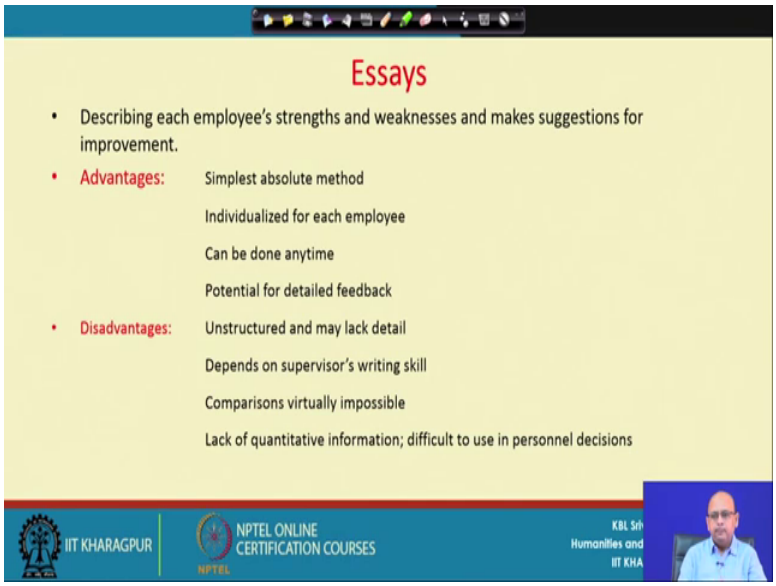
(Refer Slide Time: 24:17)



Then behavioral checklist; another absolute method, right; so you have a statements are behaviors against with the performances evaluated.

So, you have various competencies, and supervisor what he is does he simply checks it with that this behavior is shown by you are not, right. But here you are not looking at the level at which is being done. So, the advantage is that you feel impersonal and disconnected, because you do not have a scale point to use and you do not have detailed information; you simply say whether this behavior is shown or not shown by the employees, right; so in that way behavioral checklist or not that effective.

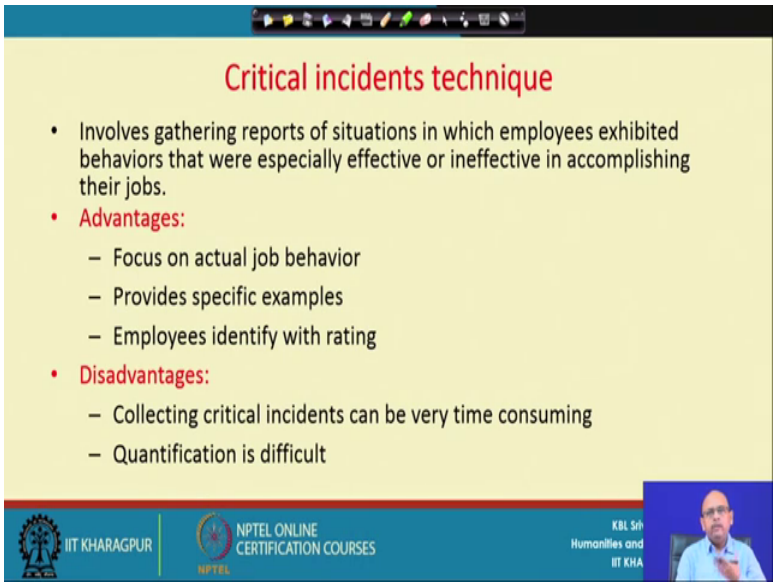
(Refer Slide Time: 24:52)



Essay is the most popular and the simple way to do the things, right. In essay what happens you simply describe his strengths and weaknesses. You know that many government organisms these kinds of confidential reports that supervisors write about their subordinates is a part of essay.

So, they simply give a description in this confidential report in the form of essay describing the strength and weaknesses of the employees and how they can improve. It also have advantage because you know these each and every employees what they have done what they have not done, and how they can be developed, but you may be biased also. How good you are whether you very good articulate and write about the employees right, suppose you have certain problems with the employees then you will not be able to write good things. So, these are the issues related to the essay.

(Refer Slide Time: 25:41)



Then critical incident: in critical incidents what happens you are going to collect information related to the employees which actually lead to very effective behavior or very ineffective behaviors.

So, you are going to focus on actual job behavior of the employees and see that what made him to perform well and what made him to perform poorly, right. So, you take specified examples and then you rate employees. The idea here is that you identify those critical incidents of behavior which contribute to performance of the employees. But one of the disadvantages of this method is that it is very difficult to collect critical incidents, right because it could be very time consuming and quantification is very difficult. Because how would quantify each and every behavior in terms of measure, because ultimately we are going to use a skill; rating a skill to measure the behavior.

So, for each critical incident whether it is very effective or not at all what you are going to do, you are going to see that how it can be made more effective and you are going to do it

Thank you very much.