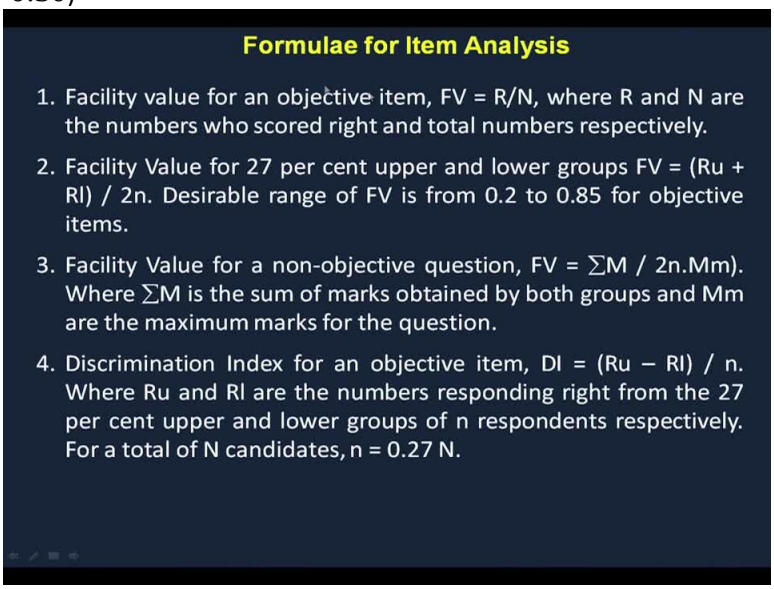
Course on Outcome based Pedagogic Principles for Effective Teaching
Professor Shyamal Kumar Das Mandal
Centre for Educational Technology
Indian Institute of Technology Kharagpur
Module 3
Lecture No 16
Lecture 16: Assessment and Evaluation (Contd)

Okay so in the last class we have discussing about that item analysis, discriminative index, facilitation value and that effectiveness of the distractors.

(Refer Slide Time: 0:30)



So in nutshell the facilitation value of an objective item is nothing but a facilitation value FV is nothing but a R by N, that means that in facilitation value if the each and each questions is getting single marks, then it is total marks R divided by N. So RN at the number whose whose score right and total number of respectively.

So suppose I have taken a test facilitation value is how easy is the test, so number of correct answer given by the upper group + number of correct answer divided by given by the lower group divided by the total number, okay and we have explained it in simple manner by giving an example. And I will share with this ppt there is a some example is already done, so you can go through it and check your things that whether it is correct or not and you can go through it.

Next is that if you see a facilitation value for that non-objective test item like that if the question papers question carries a different marks like that some question carry 5 marks, some question carry 10 marks, some question carry 20 marks like that way then also facilitation value is the total marks obtained by the upper group + total mark obtained by the lower group divided by the total question marks of that whole question set, okay.

That means that suppose there is a 10 questions and each question is 5 marks, then total marks is 10 into 5 50 marks questions, okay. And into multiply by the total group size, okay so this is where formula is given. Similarly, discriminative index the correct answer given by the upper group minus correct answer given by the lower group divided by the N, N is the group size, okay.

Now this this desirable range where discriminative index is 0.3 to 0.65 for the objective test item and facilitation value you know that you get the facilitation value. So if the facilitation value had value 0, what is the meaning? Facilitation value 0 means nobody of that nobody can answer that question. So total marks obtained by the upper group 0, total marks obtained by the lower group 0 divided by the group size is 0.

So that means if a test item has a facilitation value 0 that indicate this test item is not suitable for taking the test because either this item may be taken from taken from such a way that it is out of the syllabus of the learner or it may be not achievable but it cannot be done within this examination systems, so that is the indication. So facilitation value 0 does not have any meaning.

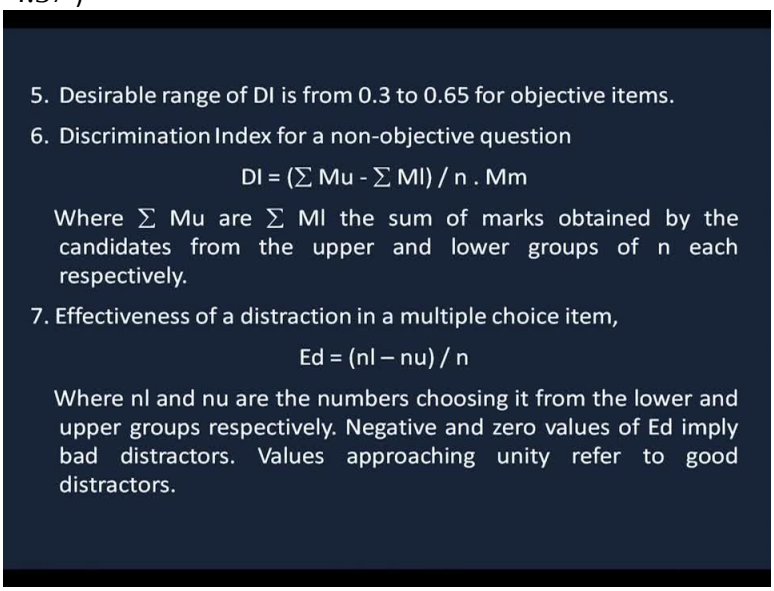
Now if I doing a non-reference summative evaluation suppose I taking a JEE examination, if a test item has fascination value 1 that means everybody can give the answer of the test item, no use of use that test item as a non-reference examination. That does not indicate that who is the good student who is the bad student. So facilitation value 0 to 1 it can vary, now if it is 0 test item is not worthy, if it is 1 then also test item is even if it is closed to 1 also not worthy.

So it is a when you design a question paper based on your test, that suppose I want to design a question paper for semester examination, then it is criteria based test. That means I have already defined my outcome of the course and I already told the students you have to achieve that

outcome I have to only test whether you achieve the outcome or not, how do I do it? I design a question paper based on the criteria which is defined by my course outcome.

If a student give that correct answer, then he (achie) achieve the outcome, if he does not then he not achieved the outcome, okay.

(Refer Slide Time: 4:37)



Similarly effectiveness of the distraction, very important very important because many people are designing the multiple choice question paper and if you see let us I can give you the example, suppose you in kids nursery level there is Olympiad exam or some other exam suppose you give a question like that a three digit number multiply by the three digit number, your objective is that student should know how to multiply a three digit number by a three digit number.

Now if you design a question paper like that 225 last digit is 5 and again the last digit is 5 of another the number. So the multiplication number also have a last digit 5, now out of four choice if you only one choice which is out last digit is 5, so the student without doing the three digit multiplication they can give the answer.

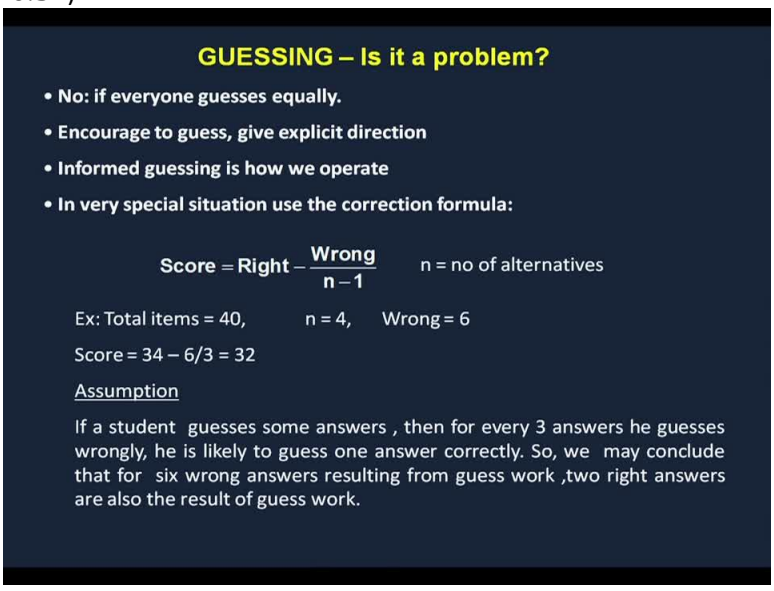
So actually objective was I want to test whether they are able to multiply three digit number by a three digit number is not tested. So effectiveness of the other distractors are not relevant, okay. So that means effectiveness of the distractor are very important and I you know that how to

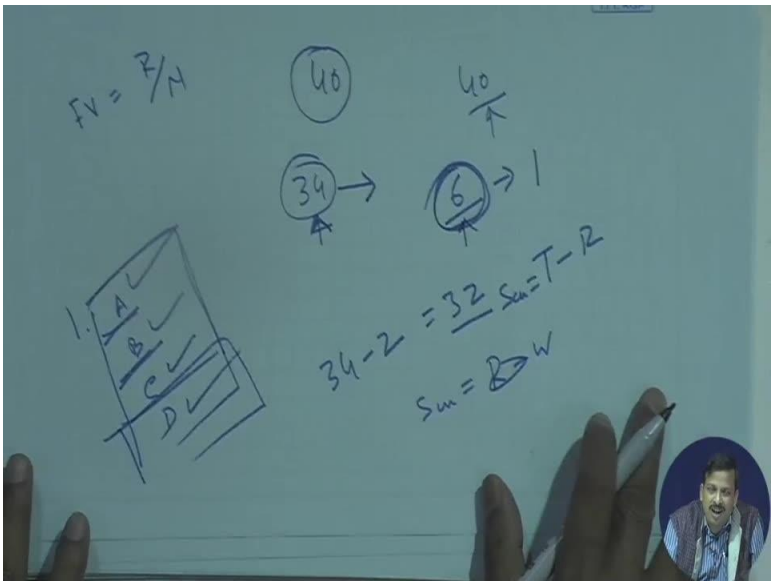
calculate the effectiveness of the distractors. And also the distractors are the valid misconception of that question paper.

So it should not be like that arbitrary distractor does not help you to test the students whether they have attain the skill or they does not attain the skill, okay. Now I therefore the another part which is called guessing, in multiple choice question paper I can either design the question paper for the negative marking that means I am discouraging the guessing, negative marking means I am discouraging the guessing, I told the student if you do not know the answer do not answer it without guessing.

But suppose in any many cases we should allow guessing, okay. So based on the requirement whether you can allow guessing or you may not allow guessing. If you allow the guessing, then how do you correct that guessing factors? There may be there may be some guess is done so you can get the correct number by the guessing.

(Refer Slide Time: 6:57)





So there is a simple solution for correction of the guessing I described and this is one way, you can do other way also. So my method is that what I am saying that that suppose I take a test for 40 marks question 40 test item and 40 marks. A student got 34 marks, I want to correct the guessing factor. Maybe student does not guess by in general I assuming student guess it, so if he got 34 marks then what would be the correct marks which is without guessing?

Okay. Now I have a four choice each test item has four choice A, B, C, D, okay so if I guess if I say my correct answer are randomly distributed over the 40 questions means I should not design a question paper where all the A are correct answer. I should not design a question paper where all the B are correct answer, where the correct answer are randomly distributed over that 40 question paper, 40 question so some question have A correct answer, some B, some C, some D, okay.

Now if a student got 34 that means, he has done mistake on 6 test item out of 40. So if he done mistake on 6 test item, what is the meaning? That if I if I have a four choice, out of three wrong choice I can get a correct choice. Out of three wrong choice that is a possibility that I can get **get** a correct choice if I tick by guessing, okay.

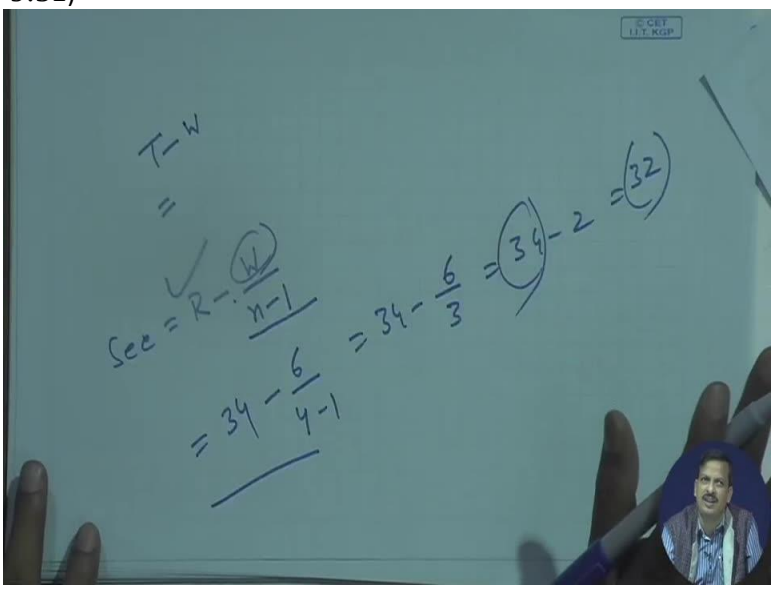
So if there is a six wrong answer, that means every three wrong answer every three wrong there is a possibility that student get one correct answer. So actual marks will be 34 minus 2 because 2

6 is wrong answer, so 2 is the guessing factor 2 marks is get by the guess, okay. So 32 marks is the correct marks without guessing, okay. So what is the general formula?

Right answer the score is nothing but the right minus wrong answer right answer minus wrong answer, okay wrong answer so score I got so 40 so 40 is the total marks. He write 34 sorry total marks minus wrong answer is the score of the student, okay.

So if I say what is the how much marks he is wrong? Wrong means how many marks he has done mistake.

(Refer Slide Time: 9:51)

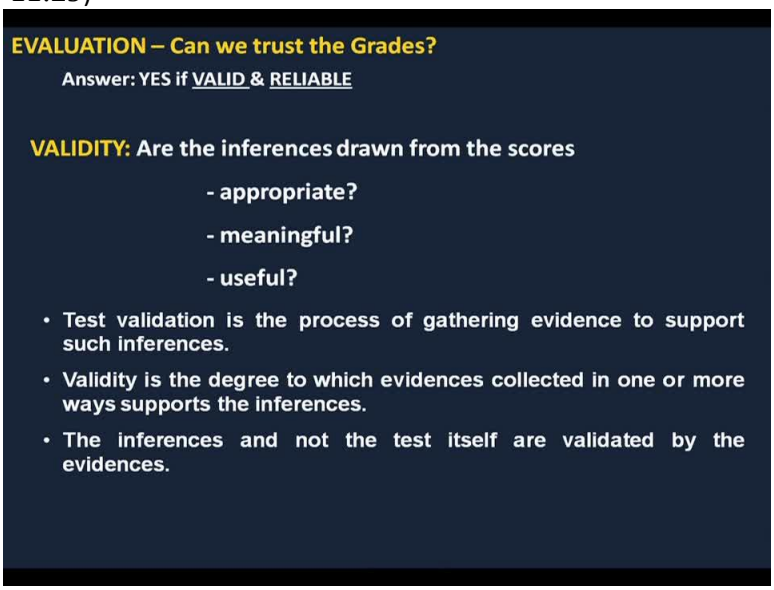


So that means total marks minus total marks he got means right answer is equal to the total mark the number he may be wrong. So if I say without guessing what should be the his number? So score will be right answer right score whatever he get minus wrong score how many number he get by wrong number divided by the number of choice minus 1.

So if say 40 is my total score, so he get 34, he got 34 now how many wrong? 6 marks is wrong divided by how many choice? 4 minus 1 is equal to 34 minus 6 by 3 is equal to 34 minus 2 is equal to 32. So right score minus wrong score divided by choice minus 1 number of choice minus 1, okay. That is the guessing factor, okay.

So if a 40 marks question paper is designed, if somebody get 34 that means he is supposed to actually get 32. You may allow guessing, you may not allow guessing that depends on your choice but if you allow guessing you can correct the guessing factor, okay.

(Refer Slide Time: 11:29)



Next one is that, yes now I know the item analysis but yes item analysis only valid I can do item analysis if I do suppose take the examination today and then I done the item analysis of this this examination. So when I design the next question paper, if that such kind of item analysis is done for thousands of questions, then probably design a question paper will be very easy to me.

Because if I tagged each and every item analysis after item analysis is tagged the score, every question paper there facilitation value, there discriminative index and effectiveness even multiple choice effectiveness distractor all are stored in my question bank. Now if I want to design a question paper let us it should be discriminative should be I require a test item whose discriminative should be for like this to this, I can pick up I require a test item whose facilitation value is like this and this, I can pick up, okay.

But when you doing a summative assessment for the semester examination, the criteria based examination there is nothing do with that facilitation value and discriminative index because I want my in my class all student should get 100 by 100 because if all are attain that skill. So my question paper is only for testing whether they are achieved the intended outcome or not.

Thumb rule is that do not provide a test item which is already discussed in the class. If you provide a test item which is already discuss in the class, then it is not that skill, it is nothing but a remembering level. Suppose you taught a design of something and you give the same design in the examination system, actually you are testing the memory of the students, not whether the students is able to design that things or not.

So do not provide a set of question paper whose all test items will be new but using the same concept. So that means I should not provide a test item for testing the competence of the two student which I have already solved. Because if I already solved the same test item, then the student will be just remember and deliver like that if you see history question paper of our higher secondary level kind of the this time the shahenshah has come, next time so Babar will come, everybody know the next time I have to write the Babar.

So everybody just mug up Babar has come in the examination center and examination hall and give that. So it is not you should test the concept of the student, okay. Then I go to the evaluation validity, suppose I design an evaluation is this evaluation is valid?

That means that can we trust the grade? valid evaluation validity means that suppose I provide a x grade to a student, can I trust that his skill is equivalent to x grade or if I take today a test more or less we will get the x grade, that is the validity. So validity means once I assign a grade to the student I evaluate the students, is my evaluation is valid?

If I say this is correct, this is the skill set of the students and yes the students acquire that skill set that is important, the validity of the skill set. So how do you design the validity? How do you test the validity? Suppose like that that half the inference drawn from the appropriate, meaningful, useful. Suppose I taught in the class AC machine design some kind of AC machine design given specification, this complexity of AC machine design I taught.

Now in question paper I solved a problem in the class and I give the same paper in the test. So a student get x grade, are he validate his the test is validate? nowhere. The next time if I test with a another set of design, he may not be able to answer. So my test is not validate, I am not saying that (())(16:03) exam of that is another process. I am saying the student is honestly giving the exam but my question paper is such that my test is fail to validity.

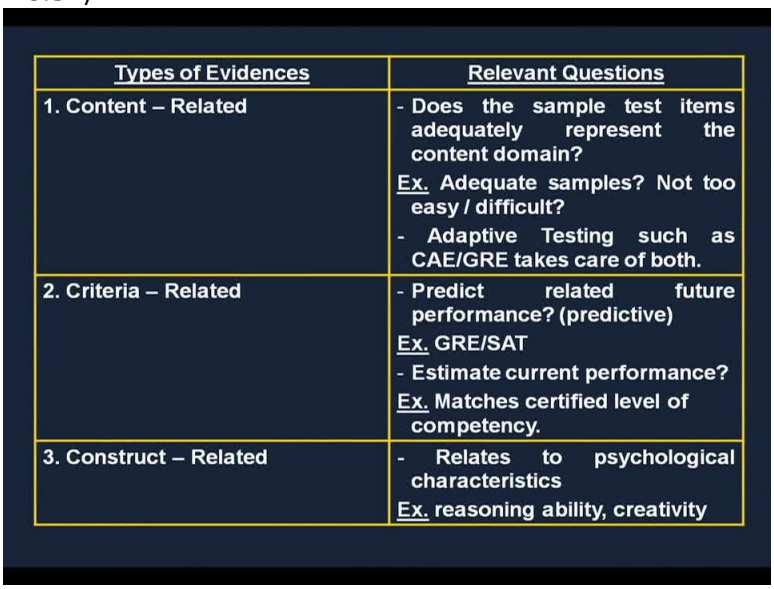
So it should be appropriate, meaningful, useful. That means suppose I taught a mathematics subject differential calculus and I taught the solving of differential equation of second order in the class. Now I design a question paper to solve the third order differential equation, none of the student can give the answer, so my test validity is not there. Similarly this is not appropriate, now suppose for a BE engineer I taught a subject which is supposed to be analysis, synthesis, application and evaluation level and my whole question paper is write a short note, define this, explain this, draw the figure, then what is the validity of this test?

Student may got the x but he does not know the application of the subject, he does not have the skill in higher cognitive level application, evaluation, synthesis and analysis level which is desired for a BE engineer. So that means if I design a question paper which is only the knowledge and comprehension level and I taught that this course is meant for that BE engineer course which made the NBA guideline whole things is wasted.

Because the student who has got the ex marks he only know the definition and derivation. So write a short note in a engineering question paper is not a valid choice, okay. So validity will be not there, so your question must be appropriate, meaningful, I will describe meaningful later on also. You should not design the question paper such that it is not no meaning of that question, it should be meaningful means it should indicate the criteria outcome of your course then it should be useful.

I should not taught the definition entropy only in case of BE engineer. If I ask in summative evaluation what is the definition of entropy? That is not that BE engineer examination, that may be ITI or like that. So it has to be useful for that group of students, so useful should be there. So validity can be inferred from the different inference if you say content does the sample test item adequately represent the content domain.

(Refer Slide Time: 18:32)



There is another scenario is there, if you see many university or many colleges or many institution has their system that out of 8, answer any five. If I remember in math exam of higher secondary examination system there is a syllabus syllabus permutation combination then there is a call Taylor series expansion all are there in the syllabus, but beginning their teacher said without touching the permutation combination and Taylor series it is possible to get the 100 marks.

Why? Because question paper is designed out of 8 you have to answer any 5. So that means if 8 represent the 100 percent coverage of the syllabus, so 5 represent 5 by 8 percent coverage of the syllabus. That means almost 30 percent of the syllabus I do not know. Still I score 100 marks. That means the validity of the question paper is gone. Then if you say okay your pass mark is 35 percent. So 70 percent syllabus 35 percent I know, that means eventually I know only 21 percent of the syllabus roughly and I passed the exam, what is the validity of that test?

Now (())(20:01) because I do not know anything about the subject but I pass the examination. So the examination (())(20:08) does not set the validity. So the does the sample test item adequately represent the content domain? That means yes I can give any 5 you can answer any 5 out of 8 no problem. But any 5 you choose, you have to cover 100 percent of the syllabus. T hat means I have to design the question paper such a way that any 5 student can choose you have to cover that 100 percent of the syllabus or 100 percent of the criteria.

That means this is criteria related. It should not be just take a test it should be criteria related, I want to test these criteria so this test item. I want to test this criteria, this test item. So it should be criteria based. Like that if I say my course objective is apply, design, evaluate, synthesis and I design a question paper describe this, define this, explain this, write a short note, so none of the criteria of my course is match with this examination system.

So that means the validity of that examination is gone. Then construct related to the that is characteristics physiological characteristics of that is there. Reasoning ability, creativity they should have applied their knowledge, not that remember and deliver. I should make a question paper they should think and apply so conceptualization of engineering problem in case of Btech engineer.

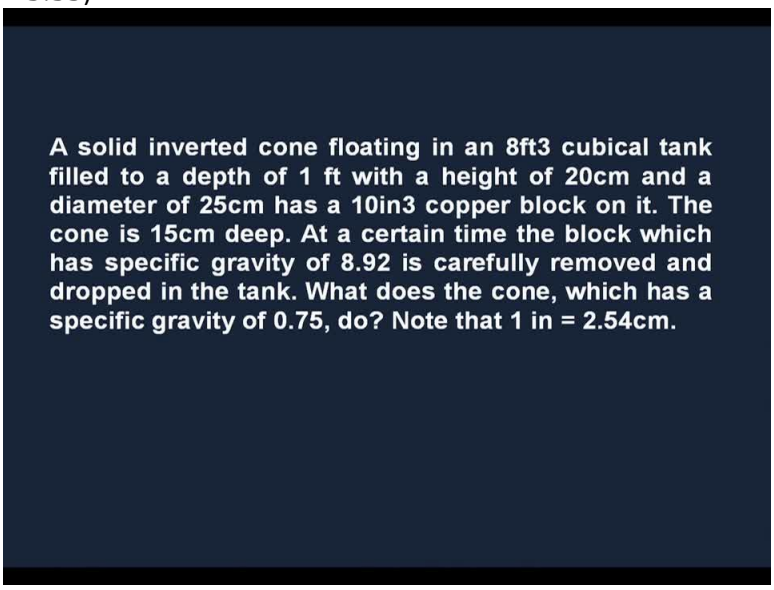
So they should think my question paper should apply a test item which they have to think, apply their knowledge to solve that problem. Prove fx equal to fy, I said it is little bit of a kind of a not validate test fx equal to proof that this equal to this, there may be the derivation is available in the book and he remember and deliver it. What is the meaning of that test item? Nothing.

I am not testing his skill. So instead of doing that I should apply I should I should design a test item which they have to think and which can they have to eager or they have to acquire that higher cognitive level skills instead of just remembering knowledge, this knowledge is not wisdom remembering knowledge, okay. So this is the validity of the test paper.

Then reliability. Again suppose today I take a test a student get 100 marks. If I design a equivalent question paper and I take the test one month after, are the student get the not the 100 marks even if 80 marks. If it is not, then what is the reliability of the test? If it is skill, suppose I acquire that skill, skill cannot be forgettable. So if I today I take a test, tomorrow I take a test pre test post test whatever I take it should be reliable over the test.

So if I design the question paper only test the skill, then it will be reliable test, okay.

(Refer Slide Time: 23:35)

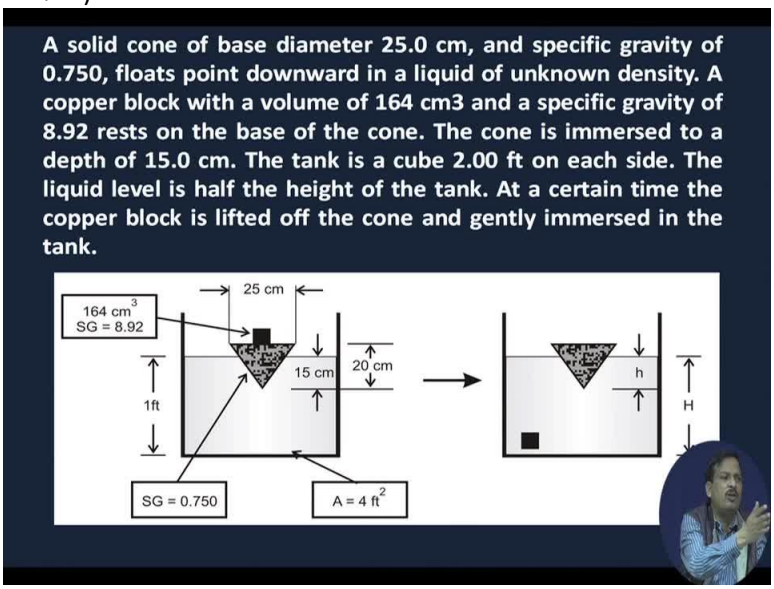


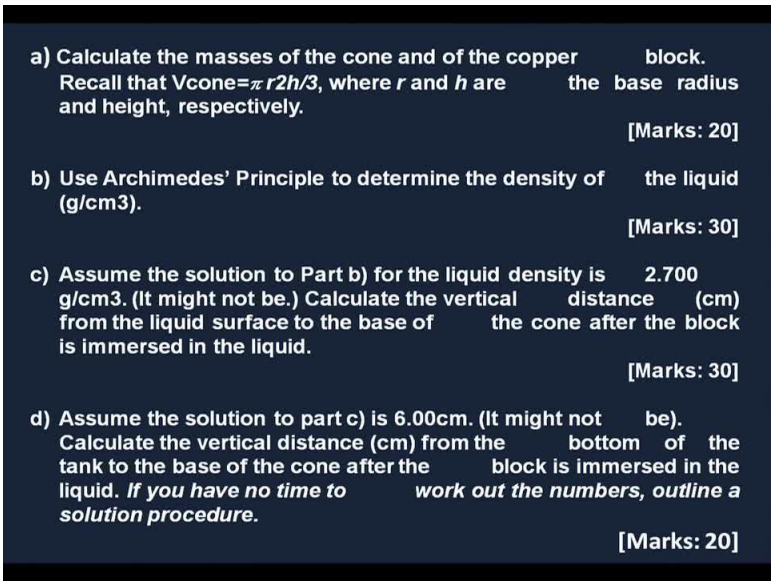
Then there is another point, suppose this question. This question taken from a some book or some internet example suppose this is the questions, if you write your question like that way and student is totally confused, what has to be done? So purpose of the evaluation system is to find out whether student attain that skill or not. Purpose of the evaluation system is not to confuse the student.

As a teacher I am very much clear I want to test this skill set so my test item should clearly mention that skill set, testing of that skill set. It is not that I give a very long questions and see what does the cone do? What is the meaning? Nothing. Suppose I want to know whether the students is know how to calculate the density, how to calculate the volume, how to calculate that kind of things.

But if you say that if I give this test item I cannot figure out whether they know this, whether they know this I know this or know this.

(Refer Slide Time: 24:47)





So instead of giving this, if I design the question like that way and make the question like this way calculate this principle to determine this, identify this then I know what exactly I want from the student. And my problem statement should very clear what has to be do. I should not I may not be provide this block diagram I may provide this text, student has to draw this block diagram and answer one by one question.

So every test item is indicated one type of skill. Later on you can do that suppose there is a multiple step of a problem, now suppose a student does not have that skill but next skill student

has. So if you he cannot do this problem this problem is linked to this problem so the value of answer of the problem 1 will be used in the answer of the problem 2.

And the student does not know the answer of the 1, then my test item should be designed such that okay if your answer if your assume that answer of the first question is within this range, do the next item. So I can specifically I can identify which skill student does not have. So purpose of the evaluation system is to test whether student has the skill or not, not to confuse the student and not to provide the marks to the student.

If you design a question paper write this, define this, square this, write a short note on this, useless that does not mean to the skill for engineering education. So when you design a question paper for BE engineer it should design that your course objective as per the (())(26:43) to the course. So that has to be tested, that is the test item design and test item analysis and evaluation and assessment.

Maybe suppose you are doing a that doing a evaluation for a seminar for a project what is doing? That, okay seminar somebody giving a seminar, somebody is sitting there and giving a marks. Student does not know which area teacher are marking him, which area he has a less confidence, which area he has a right confidence.

So before you take the presentation, you define the rubrics. Okay your presentation style will be this marks, your slide content flow will be this mark, this will be this mark so criteria you define and give it to the student before you take an assessment, then take the assessment and provide the marks and give the feedback to the students that you does not have this skill this skill this skill, now this skill you required to improve, so specifically you have to mention which skill he is required to improve.

So evaluation is not only purpose for just providing a marks, it has to tell the students yes you give the exam but your this skill and this skill is not present, so he will be acquire that skill. Now whether he acquire or not it is student responsibility, learner responsibility learner will done. But any evaluation criteria must be well defined, criteria for evaluation project evaluation you have to define the criteria okay you have to do this this marks, you have to do this this marks, you have to do this this marks, total this marks.

So I can when I take the evaluation I can tick that this marks this skill does not, this skill has, this skill has, so I can provide the marks so instead of seeing the face of the students and I use most of the lab exam we take that given a report and we ask the student to give a viva and ask a question and see the face of the student you put a marks. That does not evaluate the student, again give a marks but it does not evaluate the student.

Student does not know anything if you see most of the time lab report are same from year after year. And if you ask them they does not know anything what is the problem? Problem is that as per the (())(29:09) course or NBA guidelines student has to know how to design the experiment, how to interpret the data, how to analyze the data, how to collect the data all skill I have to test all and individual skill and provide the marks.

Then I can say yes my course design is fulfill the evaluation system. So I design a course like that I have a syllabus huge tough syllabus is available but nothing is taught, what is the use of the syllabus? Nothing. So syllabus does not improve the quality of the student, so I can say I have designed the skill, defined the skill and my examination system is only okay define this, describe this, draw this, derive this, finish then I am not testing the skill.

So examination system is not valid, examination system is not reliable, test items are not valid they are not reliable, they are not criteria based all kind of things are there, okay. So that is was about the assessment and evaluation technique, okay thank you.