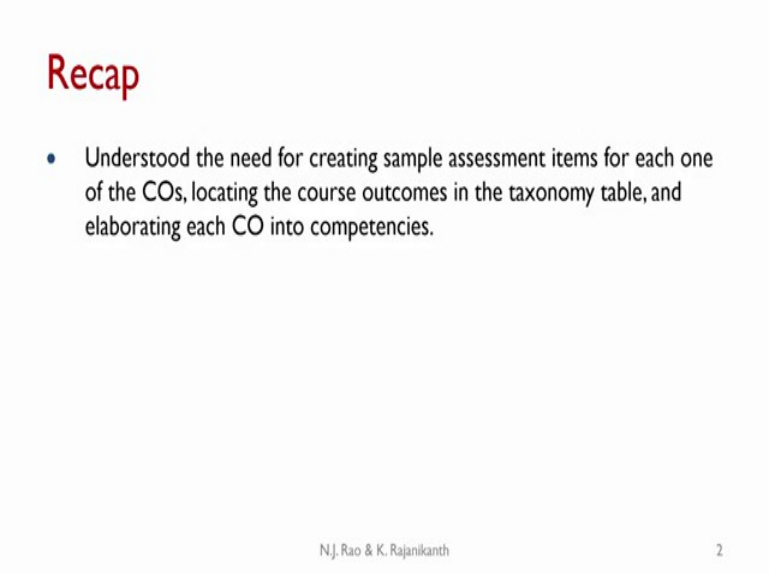
TALE - 2 Course Design and Instruction of Engineering Courses

Prof. K Rajanikanth
Former Principal - MSRIT
Indian Institute of Science, Bengaluru

Lecture - 06
Design Phase

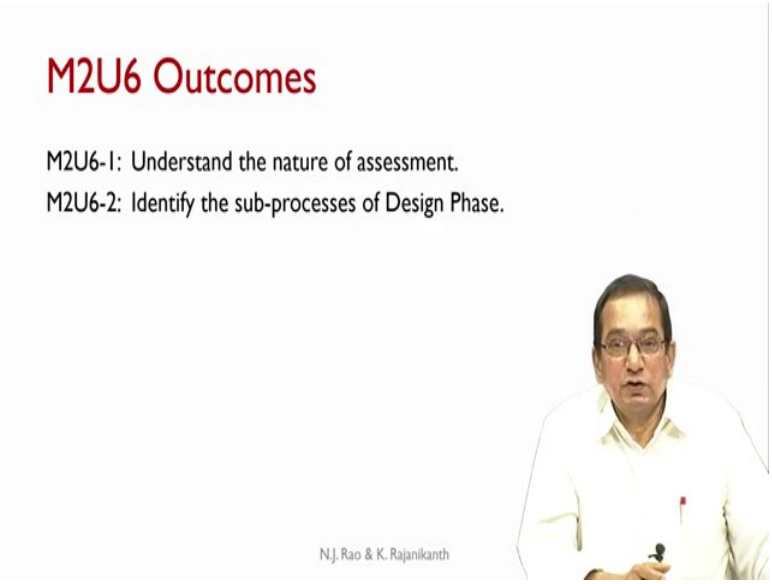
Greetings, welcome to Module 2 Unit 6 of TALE, which is on Design Phase.

(Refer Slide Time: 00:37)



We have completed the analysis phase; we understood the need for creating sample assessment items for each one of the COs, locating the course outcomes in the taxonomy table, and elaborating each CO into competencies wherever required.

(Refer Slide Time: 00:59)



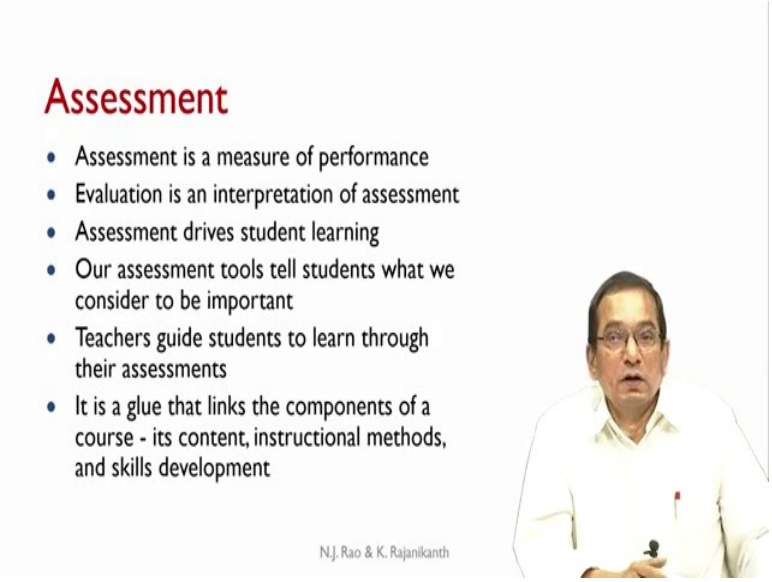
In this unit we will start with the design process of the ADDIE model. We look into the sub-processes; what is the focus of the design phase in terms of course design for an engineering program. The outcomes for this unit are: understand the nature of assessment; identify the sub-processes of design phase.

(Refer Slide Time: 01:23)

The design phase is primarily concerned with: Identifying what the learners should be able to perform to attain the learning objectives (course outcomes in the context of OBE

and NBA); and identifying acceptable levels of performance (setting the targets of attainment of course outcomes in the context of NBA.)

(Refer Slide Time: 01:55)

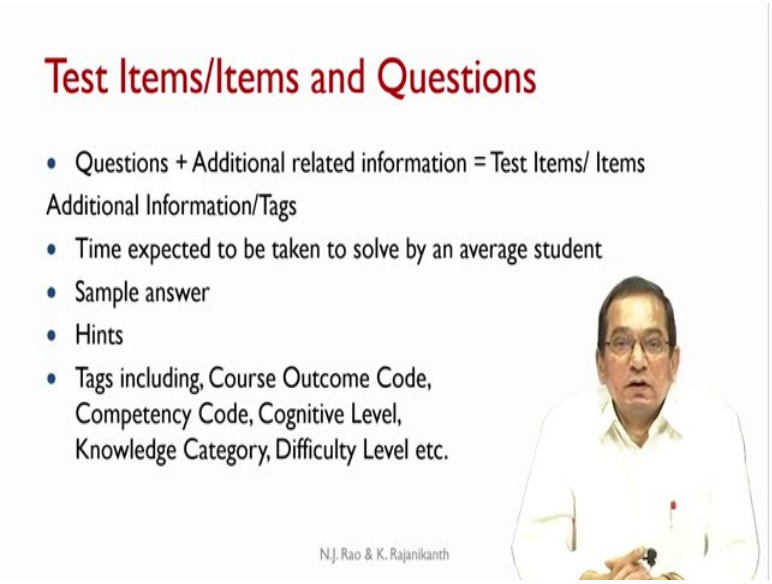


Assessment actually is a measure of performance and evaluation is an interpretation of assessment. It is true that quite often faculty uses these two terms interchangeably, but assessment is actually a measure and evaluation is an interpretation. It is very important to see that assessment drives student learning. Many have a notion that assessment is really not that important, but it is a known fact that the quality of assessment drives the quality of learning.

Assessment is a high stake item for the students, so it is a source of tension; but it is also the only means the instructor has to drive quality learning in the students. Our assessment tools tell the students what we consider to be important. Teachers guide the students to learn through their assessments. So, it is absolutely essential that the assessment is of high quality and it should help the students learn deeply and it should guide the students into learning deeply.

Assessment plays a crucial role. In a way it is the glue that links the components of a course - the contents of the course, the instructional methods and the skills development. These are all glued together through assessment. A good quality assessment is essential to ensure good quality learning.

(Refer Slide Time: 03:47)

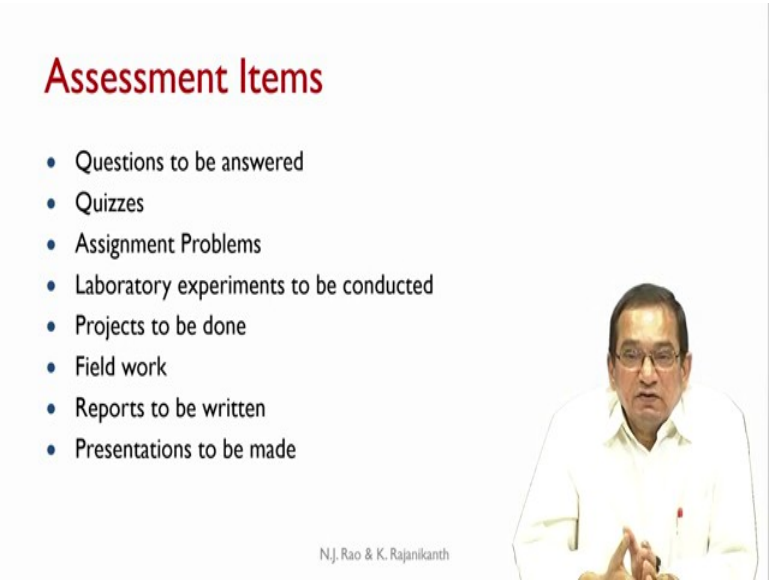


What do Test Item or Items and Questions have as the relationship? Questions plus additional related information is generally called as a test item or item. What additional information or tags we can provide with the questions? Time expected to be taken to solve by an average student. This is very essential to ensure that the total question paper or assessment instrument as technically it is called has the questions or items whose total time is reasonable - in the sense that the average student should be able to answer the required number of questions comfortably.

It is necessary to have an estimate of the time expected to be taken to solve by an average student. This is definitely an estimate by the instructor, but it does help in planning the assessment instrument properly. We can also provide a sample answer as a guideline for ourselves as well as for evaluators if they are different from the paper setter; certain hints can be given. We can also include tags - course outcome code, competency code, cognitive level, knowledge category, difficulty level etcetera.

When we create an item bank - as we shall see in a later unit - with these tags it becomes relatively simpler to set an assessment instrument, by drawing on the assessment items which are relevant for that particular context.

(Refer Slide Time: 05:39)

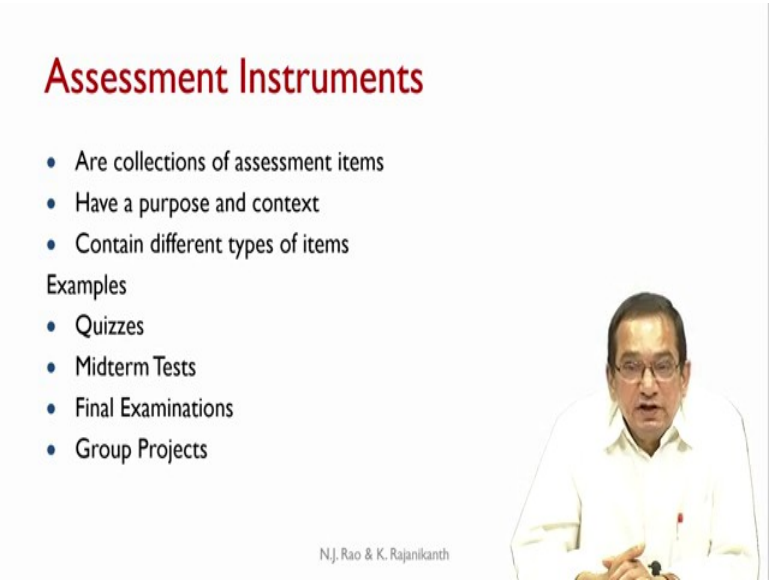


What kinds of Assessment Items are possible? A wide variety of assessment items are possible - questions to be answered, quizzes we can have; it is also important to note that a quiz need not contain only items at remember level though often that is the case! It is possible to have a quiz containing items which are at understand level also, probably apply level with lower difficulties. Higher levels may be difficult, but it is possible to have a quiz containing items of different cognitive levels.

We can have assignment problems. Often assignment problems should be focusing on higher cognitive levels. Because this is a take home assignment, students get reasonable time to complete the assignment. What we will not be able to cover in fixed time assessment instruments like class test, we can cover them in an assignment. Still instructor has the freedom. So, if he or she wishes certain number of questions at the remember level, at understand level, can also be included in assignment though it is much more common to have higher cognitive levels in an assignment.

Laboratory experiments to be conducted, projects to be done - projects could be even small, mini, group projects or it can be a project of major importance. Fieldwork is possible. Reports to be written, presentations to be made - all these come under assessment items. The selection of appropriate assessment items is the prerogative of the instructor.

(Refer Slide Time: 07:29)

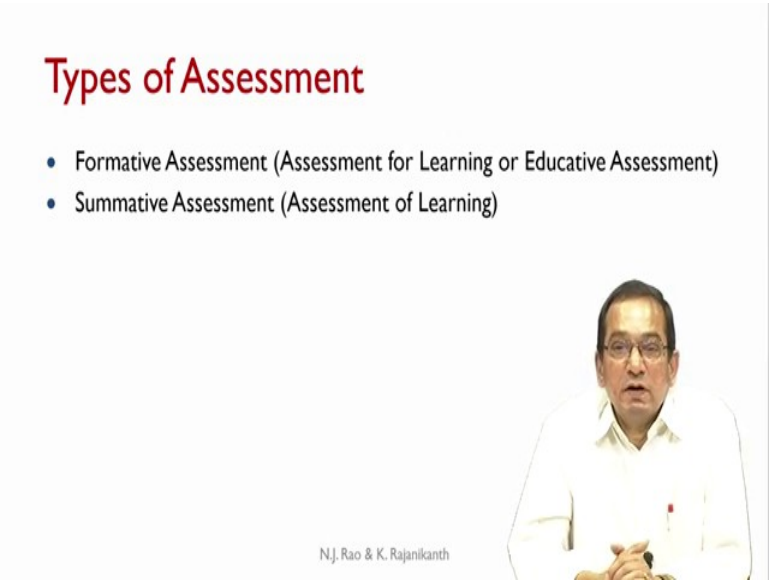


An assessment instrument essentially is a collection of assessment items. The instructor selects certain assessment items, combines them to form the assessment instrument. Typically, a question paper if it is a written performance. But it is possible to have an assessment instrument with other assessment items also. I could create a question paper with quizzes and questions which require written responses. So, it is possible to have diverse range of assessment instruments; instructor has to decide on which are the assessment instruments that he or she should use.

There are guidelines possibly available either from the institute or from the university as to the nature of assessment instruments allowed, but within those guidelines instructor would be having certain amount of freedom in choosing the assessment instruments. So, they have a purpose and a context and that will determine the structure of the assessment instrument. They can contain different types of items. Examples are quizzes, midterm tests, final examinations, group projects.

As mentioned there are guidelines provided either by the institute or by the university with respect to particularly items like midterm test and final examinations. But, within these guidelines certain freedom is available to the instructor to choose the items, which constitute the assessment instrument.

(Refer Slide Time: 09:05)



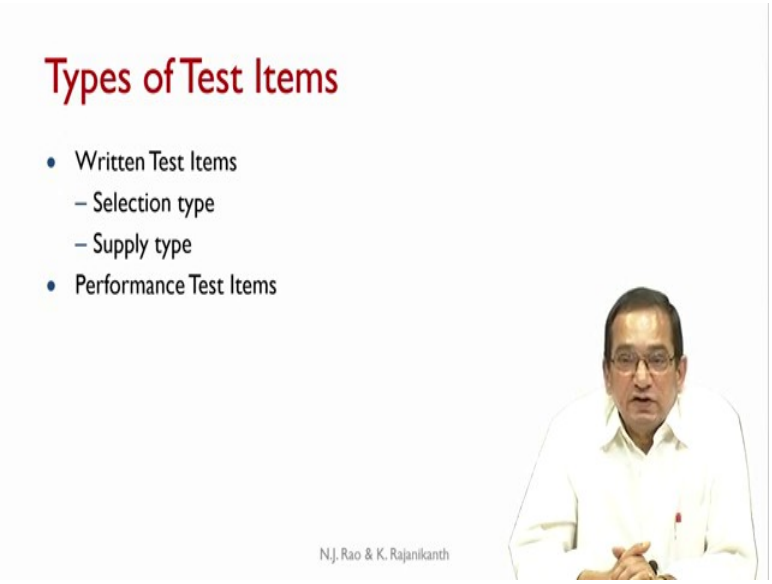
Types of Assessment: broadly we have formative assessment or assessment for learning or educative assessment - this is also called as diagnostic assessment in some cases; and summative assessment where assessment of learning happens. The purpose of a summative assessment is to finally give mark or a grade to the student. So, it is summative and it has a final grade or marks associated with that particular assessment.

In formative assessment the interest is more on determining whether the learning is happening the way intended by the instructor, or are there any bottlenecks in learning, or any sticky points in learning which require some kind of a mid-course correction, some kind of a change of the instructional strategies. The instructor is actually using the assessment as a tool to see if learning is happening in the intended way; that is why it is assessment for learning or a educative assessment.

These results are generally used to effect a change in the instructional methods or strategies, if dictated by the results of the formative assessment. It is possible that a particular instrument is used both as a formative instrument as well as summative instrument.

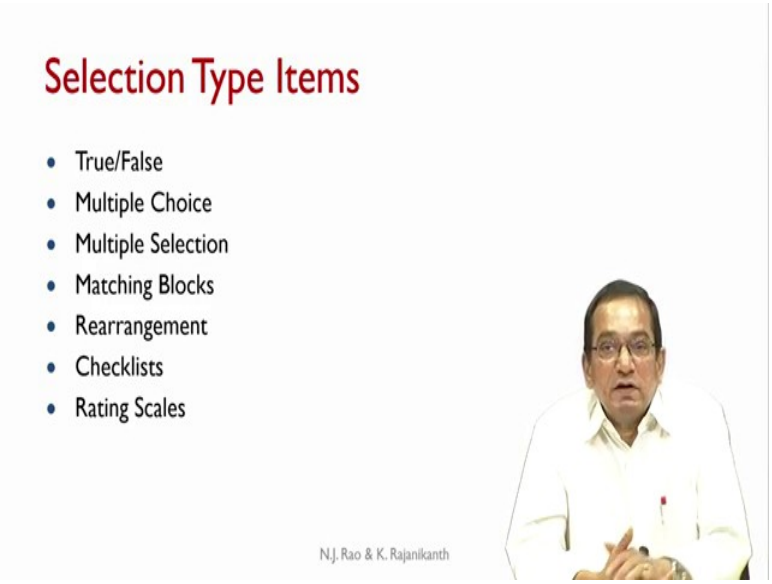
For example, a class test could be used as a summative instrument, what happens typically in current context, but it also could be used as a part of the formative assessment by trying to look at the responses given by the students to determine, if there are any impediments to quality learning.

(Refer Slide Time: 10:55)



The broad classification of the test item would be that, there are written test items where the performance of the student is actually in the form of a written submission or performance test items. So, the performance test item can be something like a presentation, where there is a performance and that needs to be evaluated to some extent in some subjective fashion; we look at that presently.

(Refer Slide Time: 11:21)



The first type where it is a written response type; again you have wide choice of items possible. Primarily they can be called as selection type or supply type. In the selection

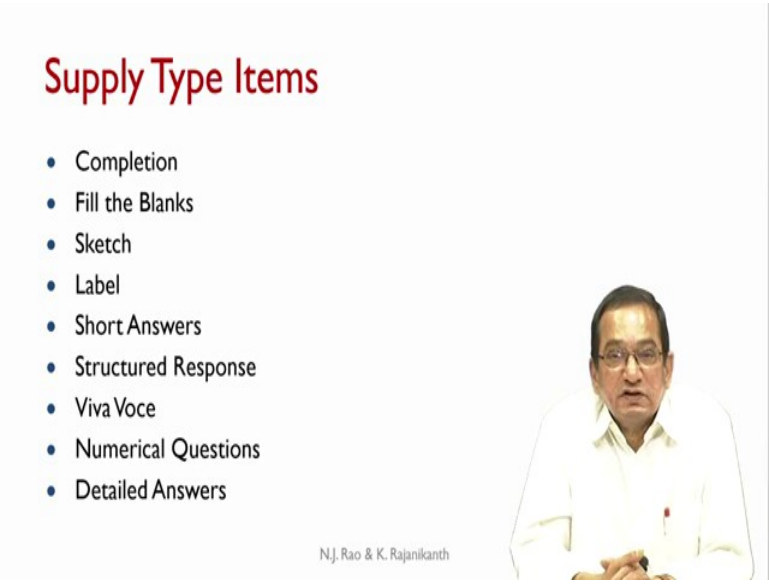
type the student is presented with certain responses and the student selects one or more of these as the response to be given to the question. A typical example would be a multiple choice question.

In the supply type the student has to supply the answer, based on his or her understanding or long term memory of the apply procedure or other kinds of a clues. Basically the selection type again has a wide range. Primarily the typical format is like multiple choice, but we could also have multiple selections, where there is more than one right answer, or true or false statements, or matching blocks.

Again in the matching blocks we have a choice - both sides can have same number of items or different number of items. Re-arrangement: particularly a sequence is given and the sequence contains the right steps, but in wrong order. Student is expected to reorder them in the proper fashion; the criteria for reordering could be the underlying logic or the underlying temporal sequence.

This is one possible way. Again it depends upon the content, the nature of the course, the nature of the assessment instrument and the context and where something is more suitable than some other kind of assessment item. Instructor has to decide which are all the items best suited for the given context. There can be checklists; there can be rating scales; a wide variety is possible definitely. Even in fill up the blanks, fill up the blanks with a word chosen from a given list of words can become selection type.

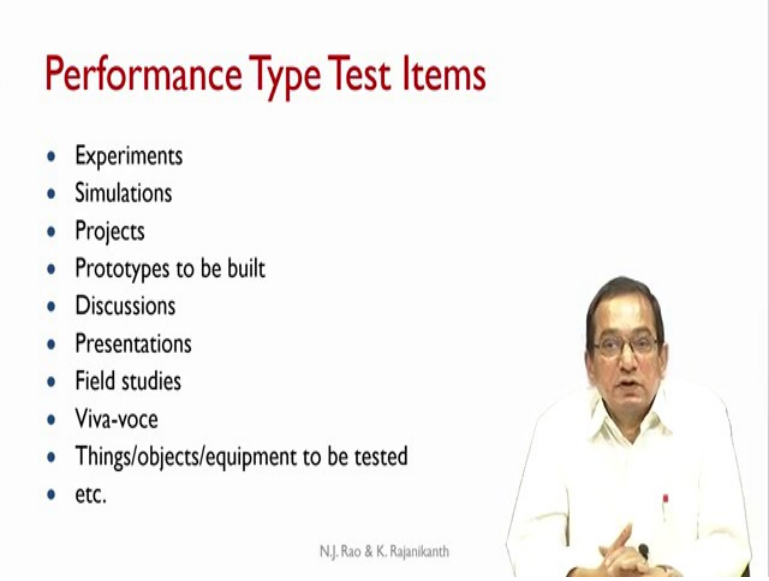
(Refer Slide Time: 13:25)



Supply Type Items are fairly well known. Our typical detailed answer questions that we ask in a test or semester end question paper. We can also have completion type, fill in the blanks where the words are chosen supplied by the student, there is no prior list which is given; or it can be sketch, label, short answers, structured responses, numerical questions, even viva voce.

Though some do feel that viva voce is more like a performance, because it is very difficult to eliminate the subjective feel for the performance aspect of viva. There is a body language, there is a way in which the answer is expressed, there is a way in which the communication takes place nonverbally - these factors are very difficult to be eliminated. So, some would consider vive voce as more of a performance type, but it is possible to consider that as a supply type item also.

(Refer Slide Time: 14:31)

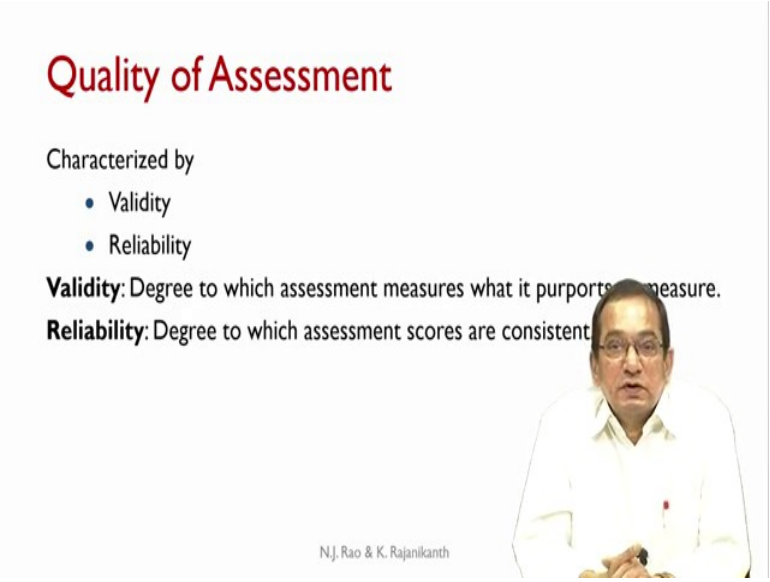


Then performance type test items are typically like presentations. A student comes onto the Dias and presents a seminar or presents a synopsis of the project. It could be group discussions, projects, simulations, experiments, and prototypes to be built and demonstrated, field studies on which some kind of presentation is made. As mentioned, viva-voce can be considered as a performance type, things/objects/equipment to be tested and a summary of the test results to be presented - there are so many varieties possible.

The primary distinguishing feature of a performance type is that there is more than merely the response being correct or not correct. So, there are other aspects to be

considered which essentially are subjective, but we need to bring in certain objectivity to the essentially subjective assessment.

(Refer Slide Time: 15:33)



Another important thing is the quality of the assessment, which is characterized by validity and reliability. Validity is the degree to which the assessment measures, what it purports to measure. The validity from a theoretical perspective can be considered in a variety of ways. But from our context the primary focus is that, the assessment instruments are being used to elicit responses from the students, which will provide us data based on which we can determine the extent to which a CO is being attained. CO is essentially a competency or a set of competencies that a student is supposed to acquire and demonstrate.

We would like to know, what is the extent to which the student has acquired these competencies and he is able to demonstrate these competencies? Hence we are using the assessment. Unless the assessment is valid, the performance data that we get from the students will not be of any help to us in determining the level of attainment of that CO. If, the CO is stating certain set of competencies, but if the assessment is really not addressing those competencies, then, the assessment instrument is really not valid; the data that we get is of no use to us in determining, what is the level to which the students have acquired those competencies.

It is very essential that every assessment instrument be checked for validity. This is very important as under the NBA framework colleges are expected to have a committee to scrutinize the assessment instruments for quality. One of the very important factors of quality is validity. The committee is supposed to actually check the assessment instrument for validity. An invalid instrument is practically useless for us in terms of getting the data. Even semester end examination papers have to go through certain amount of quality assurance activity.

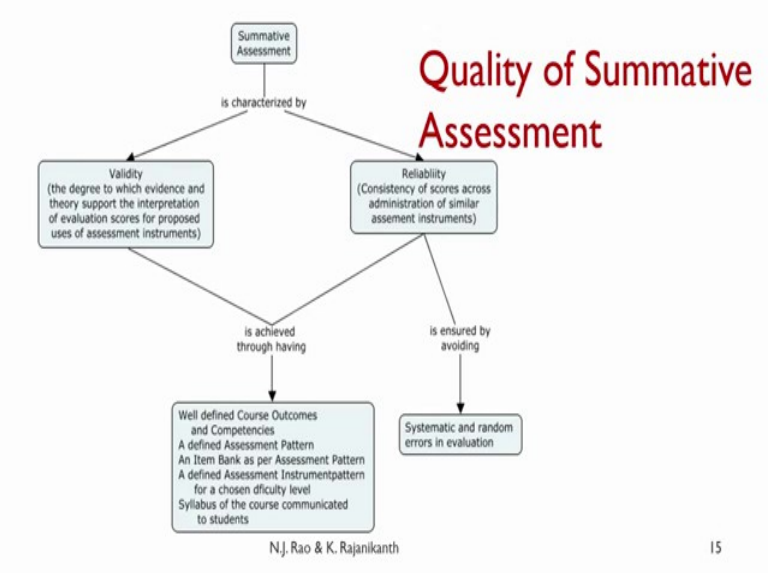
Even, if it is a tier 2 institute- that means, academically non-autonomous institute, where the semester end examination is held by the university - the university itself is supposed to have a board of examiners, which really looks into the semester and examination papers and checks the for quality. Therefore the validity is an extremely important aspect and unless the instructor exercises considerable care in advance planning and execution of that plan properly, the assessment may prove to be of not much use in terms of the OBE framework.

The second important aspect of assessment, which generally is not considered well by many faculties, is reliability - degree to which the assessment scores are consistent. If one batch of students go through a course; its assessment, we get certain performance data; and if another batch goes through the same course; assessment; and we have the performance data and if the performance data is similar, can we expect that the level of learning is also similar in both batches?

There is another way of looking at it. In one batch we have seen that the attainment levels are lower than the targets, which means we have incorporated certain action plans next time to improve the level of attainment. We do get performance data, which shows an improvement in the level of attainment. Can we say that actually the learning has improved or the improved performance data is because of poorer assessment? To what extent our assessment is reliable in the sense that scores obtained from different batches, at different times have certain consistency. They indicate data which essentially reflects the learning. That means the assessment instruments administered are more or less similar in their nature, in terms of extracting the performance of the students.

The reliability is also important, but the reliability can be assured by following certain process steps. The quality of assessment essentially would determine, how useful is the performance data for us in terms of ensuring quality learning by the students?

(Refer Slide Time: 20:37)



Quality of the summative assessment is shown in this concept map. The summative assessment is characterized by validity and reliability.

The validity is the degree to which evidence and theory support the interpretation of evaluation scores for proposed use of assessment instruments - the key word is proposed use. If we are using this data to determine the CO attainment levels, really this data must be valid for determining the CO attainment level. That is the validity of the summative assessment or a particular assessment instrument. Reliability is consistency of scores across administration of similar assessment instruments over different batches. The validity and reliability both are very important for us and both of them can be achieved through following certain process steps.

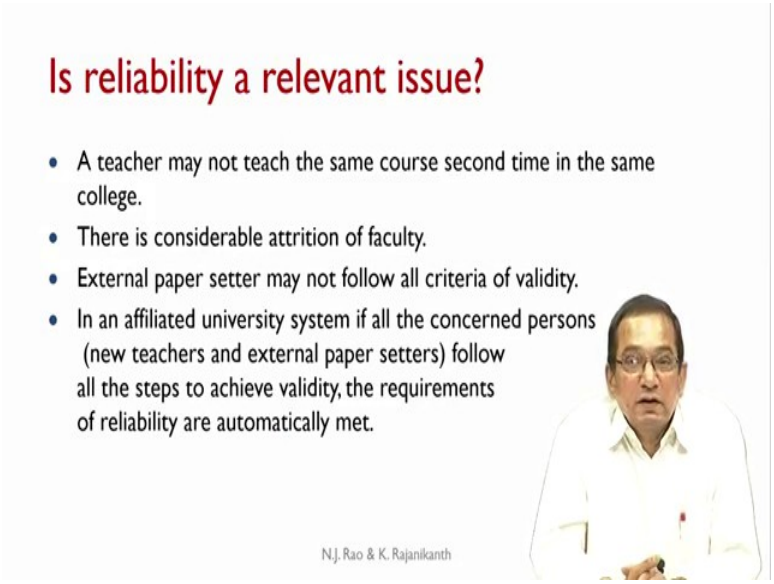
The process steps involved are: first well defined course outcomes, which can be expanded to competencies as required and these course outcomes are communicated to the students up front along with the syllabus; a well-defined assessment pattern that is essential to ensure that the assessment covers the CO properly, at proper cognitive levels, (we will discuss that again.); an item bank as per the assessment pattern, which makes it

relatively easy to create an assessment instrument that is valid and reliable; a defined assessment instrument pattern for a chosen difficulty level.

The difficulty level (we will discuss again) is different from complexity level, that is if you have got an assessment item at the level of understand and another one at the level of apply; apply is at higher complexity level, but, that is not to say that it is at higher difficulty level. At a given complexity level we can have assessment items of different difficulty levels. At the same level of apply we can have an item which is very difficult, less difficult, moderately difficult - like this different difficulty levels are possible. The instructor must decide on the difficulty levels of the assessment items that he or she is planning to use in an assessment instrument. By ensuring these process steps we can ensure the validity and reliability of the summative assessment.

The reliability is also ensured by avoiding systematic and random errors in evaluation. This again is facilitated when the evaluation can be automated to the extent possible or when the evaluation is by a faculty against well established solution patterns. Validity and reliability play a very important role in ensuring quality assessment and this requires fair amount of planning upfront. We will look at what kind of planning is required, but that is essential to ensure quality of the summative assessment.

(Refer Slide Time: 23:55)



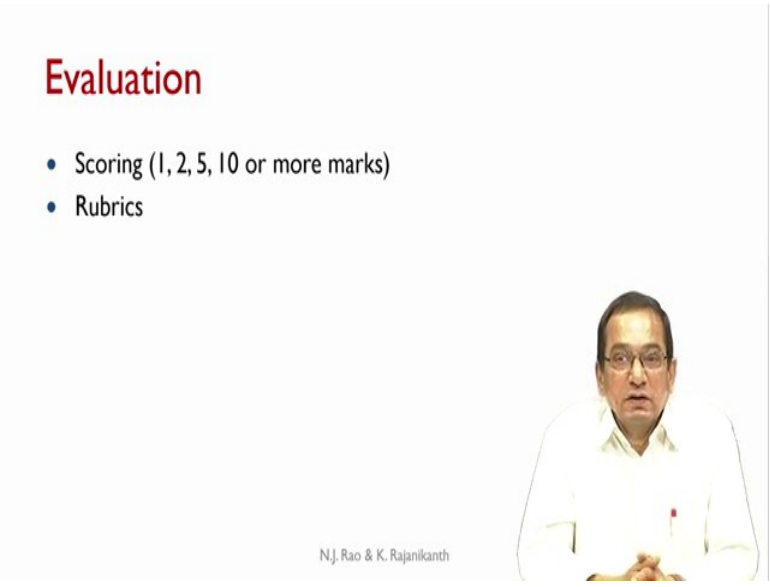
Sometimes people do wonder whether reliability is really a relevant issue, because a teacher may not teach the same course second time in the same college! There is

considerable attrition of faculty in most of the engineering colleges except a small number of higher level institutes. External paper setter may not follow all criteria of validity. In affiliated university system, if all the concerned persons - new teachers, external paper setters - follow all these steps (that is very important that we must have a process and all must follow those process steps); if that happens then validity issue and the requirements of reliability are automatically met. This essentially means that if we can set up a process, have it implemented properly and have all the relevant people - the new teachers, the faculty, the external paper centers - all of them follow the process steps properly, then we get validity as well as reliability. Hence it is important to have the process implemented properly.

It looks like some kind of imposition of external rules on the freedom of the faculty, but it is not really that. It is essentially - faculty has all the freedom - but the process steps ensure that what the faculty does results in quality assessment. It should not be seen as a kind of a straightjacket or as a kind of a restriction on the freedom of the faculty, but it should be seen as a process initiative to ensure quality of the summative assessment.

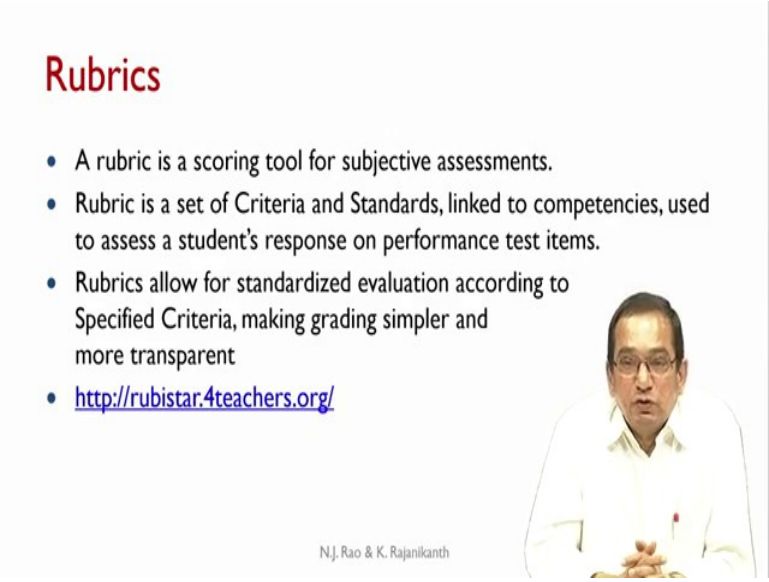
Looked at from that perspective faculty should have no problem in adopting that process step. It is possible to fine tune the process according to the rules and regulations of the institute or the university, but we must have a process.

(Refer Slide Time: 26:01)



Then evaluation is scoring. There is again fair amount of variability typically 1, 2, 5, 10 or more marks, but several semester end question papers do have different kinds of scales; even 12 marks, 8 marks - these are all possible. But, again we will see later that having certain kind of a standardization across the institute or the university has certain freedom in terms of creating a good item bank, but again this a policy matter! But if it can be taken care of, then we can create quality assessment relatively with less amount of effort. We need rubrics when we are trying to evaluate the performance oriented items.

(Refer Slide Time: 26:43)



What is rubric? It is a scoring tool for subjective assessment of a performance, like a student making a presentation. Quality of presentation is measured not only by the actual content, but there are several other factors which make the presentation a good one.

What are all those attributes which make the presentation good and then how do we measure those, how do we evaluate them; they all come under the broad antenna of rubrics. Essentially it is subjective, even with rubrics subjectivity is not altogether eliminated, but it does give an objective framework and it makes it transparent and it also creates some kind of structure, which is helpful to the student also. When the rubrics are communicated to the student, student knows what are the attributes on which he or she is being assessed. That would help the student also to understand the process of assessment more clearly and prepare accordingly.

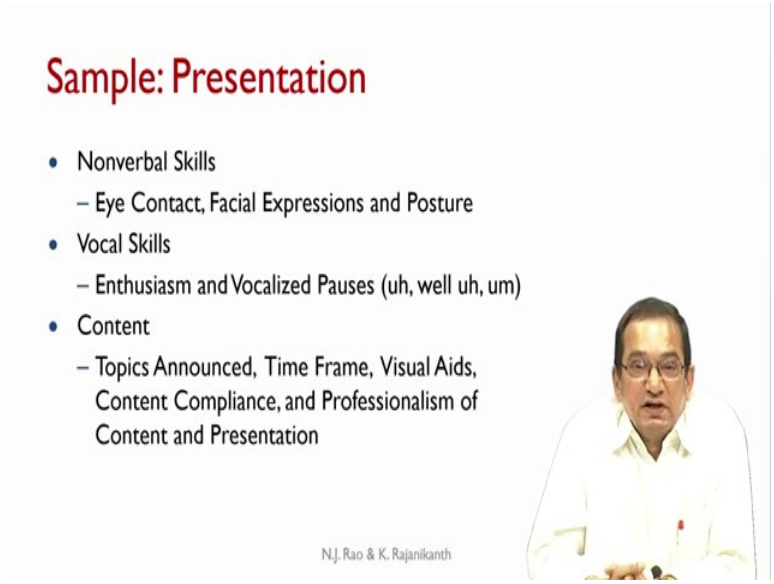
Rubrics are a set of criteria and standards, linked to the competencies, used to assess a student's response on performance test items. Typically, like a seminar presentation or a project work presentation. Rubrics allow for a standardized evaluation according to specified criteria; making grading simpler and more transparent. And when these are communicated up front to the students - that means, when the rubrics are communicated to the students - students know what is the basis on which they are being evaluated and they can prepare accordingly.

If the whole system becomes more transparent, what is essentially a subjective evaluation has now become a more transparent no doubt still subjective, but a more transparent framework based evaluation. That is the advantage of rubrics and there is a website which is mentioned there, rubistar.4teachers.org, which contains varieties of rubrics, for varieties of items, for varieties of courses. They can be used as guidelines.

Several other education centers of several universities do provide some kind of guidelines on how these rubrics can be constructed, like the learning teaching center of Carnegie Mellon University. Many other Universities do provide be kind of a guidelines based on which the rubrics can be constructed.

But, most important point of the rubrics is that they are first constructed; that means, rubrics exist and second: the rubrics are shared upfront with students. This is an essential aspect to make rubrics effective.

(Refer Slide Time: 29:33)



As an example we can look at how we evaluate the presentation. The rubrics for that could contain, primarily at the highest level, 3 categories - the content itself which typically is based on the course, but there are other parameters on which the valuation takes place; they can be vocal skills and nonverbal skills.

In nonverbal skills we could have eye contact, facial expressions, posture; there could be several other attributes that you wish to list. In vocal skills you can have the enthusiasm and vocalized pauses (uh, well, um, aah) - the kind of pauses that we make during the presentation, how frequent they are; are they impeding the progress of the flow, are they impeding the ease of communication?

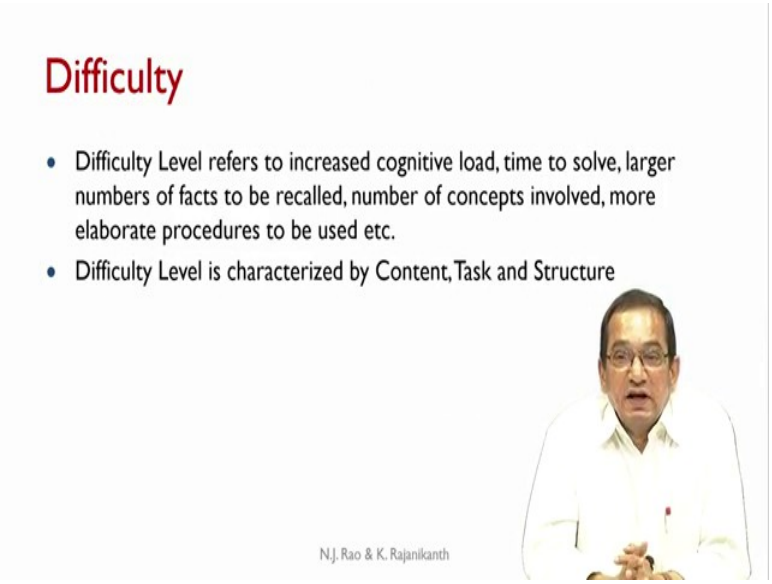
One could have parameters like filler words that people use or the completeness in terms of syntax, and semantics of the sentences used by the student etc. These attributes have to be listed by the faculty up front and communicated to the students.

The actual content itself, the topics announced, the timeframe - given certain amount of time, to what extent the timeframe is being maintained by the student, what kind of visual aids are being used, whether the content compliance is happening and professionalism of content and presentation. There can be several attributes like this and they can be organized hierarchically that is at the top of the hierarchy we have nonverbal skills, which is broken down into a lower level like eye contact, facial expression, posture, gestures, etc.

For each of the attributes listed certain levels of performance have to be determined. The performance levels can be approximately 3 to 5 because the more the performance levels it is more difficult to evaluate and cumbersome; lesser the performance levels - the evaluation will be very coarse. For each level we should describe under what conditions the student's performance is expected to be at that level.

For example, eye contact: maximum marks allocated to that particular attribute is 3; describe the performance/criteria as to under what conditions does a student get the whole 3 marks. The student can train himself or herself to ensure that this particular attribute is taken into account in his performance when this criterion is communicated. This kind of a rubrics preparation and sharing them upfront with the students would help the students do well, in terms of the performance related/oriented test items. This is a very important tool that we must have.

(Refer Slide Time: 32:41)



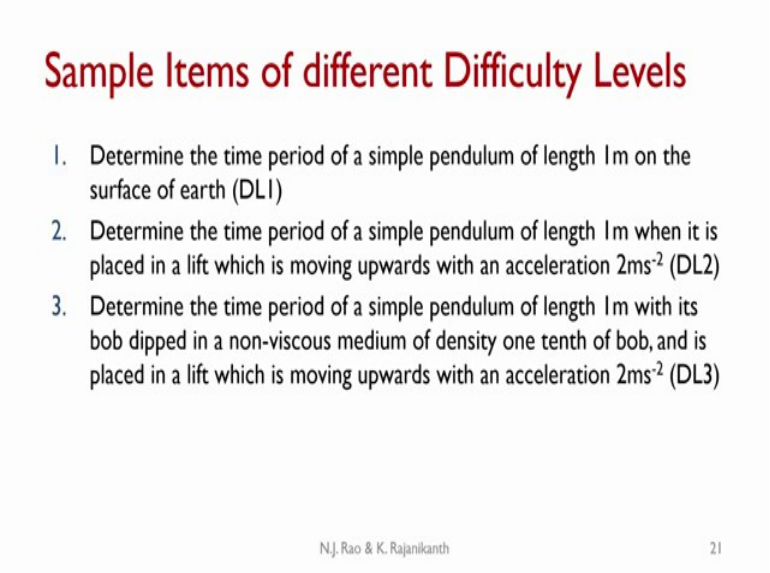
Difficulty: Another very important aspect of the assessment items. Often some of the faculty gives this as a kind of an explanation, why the assessment items that they have created are at lower complexity level! If the CO is at apply level, they may ask a question at an understand level and say that our students are weaker students, they will not be able to answer at the apply level.

This results from confusion between difficulty level and complex level. These two are two different things, the complexity refers to the cognitive level, but the difficulty level refers to the amount of effort taken by the student, the time taken to solve the problem, the cognitive load, the number of facts to be taken into account, the number of concepts to be taken into account, but at the same cognitive level. That means, at the apply cognitive level itself, I can now have assessment items of different difficulty levels. We could classify them into any number of levels, but typically people classify them into 3 levels - very low difficulty level, moderate difficulty level, high difficulty level.

At a given cognitive level, it is possible to construct assessment items at different difficulty levels. In order to do justice to the students, in order to help them learn well, it is better that we stick to the cognitive level, we do not compromise on the complexity, we stick to the cognitive level. But at that cognitive level we can compose assessment items of lower difficulty, if we feel that that is what we need from our students, but

compromising on the cognitive level itself would be a disservice to the students. Difficulty level is characterized by content, task and structure.

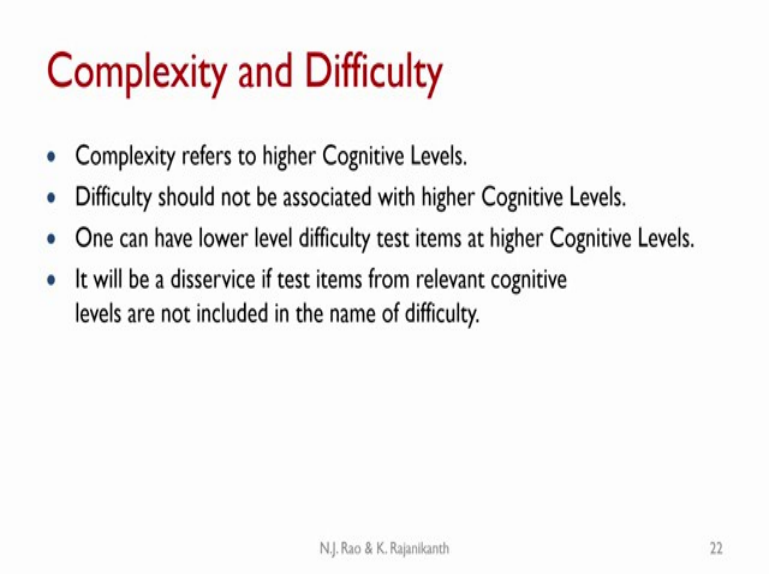
(Refer Slide Time: 34:47)



Example: at the same cognitive level questions of three different difficulty levels are posed. The first one is lower difficulty, second one is moderate difficulty, and the third one is higher difficulty level. All of them are concerned with determining the time period only, which is basically at the cognitive level of apply; the time period of a simple pendulum of length 1 meter - so that much is common. But, in the lowest difficulty level you just asked for this time period calculation on the surface of the earth.

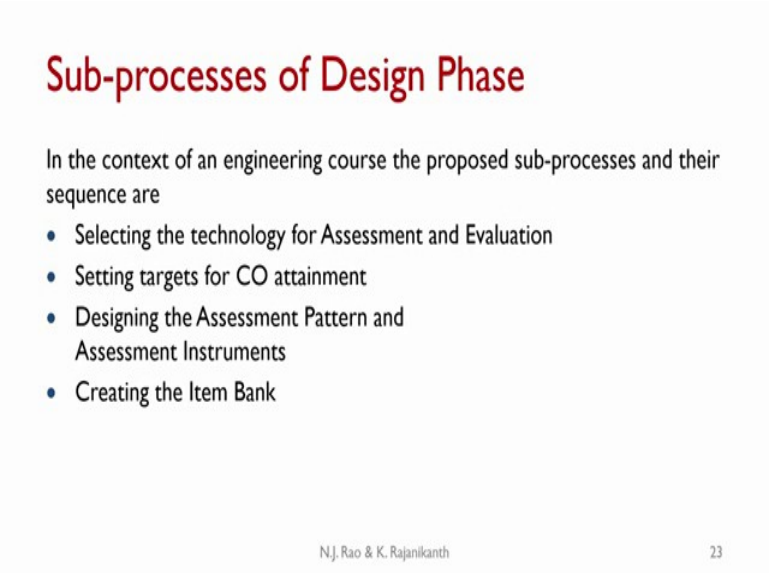
In the second one the same information is to be calculated; now the pendulum is placed in a lift, which is moving upwards with an acceleration of 2 meters per second square. It is at higher difficulty level while retaining the same cognitive level. The third one is also same - determining the time period of a simple pendulum of length 1 meter - but now the context is the bob dipped in a non-viscous medium of density one tenth of the bob, and is placed in lift which is moving upwards with an acceleration of 2 meters per second square. This is higher difficulty level than second one, but all of them are at the same cognitive level.

(Refer Slide Time: 36:05)



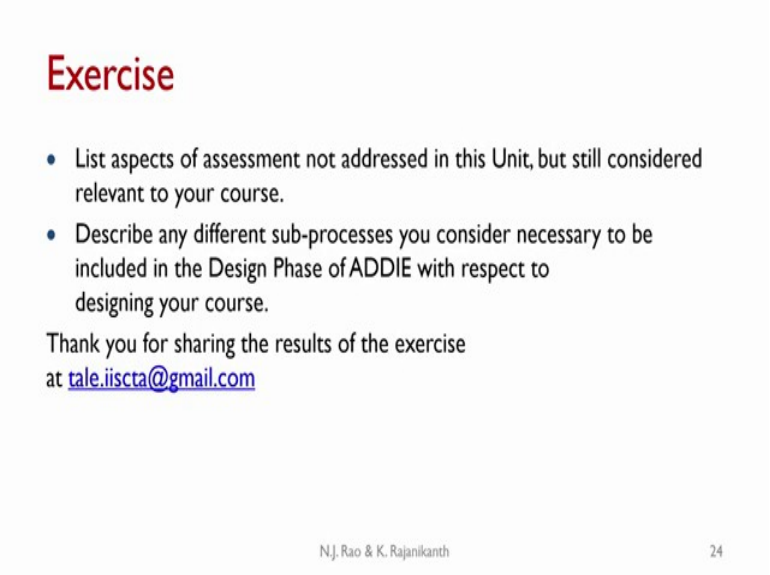
It is very important that we do not sacrifice the cognitive level, but we can tone down the question in terms of difficulty levels. Complexity refers to higher cognitive levels; difficulty should not be associated with higher cognitive levels. One can have lower level difficulty test items at higher cognitive levels! It will be a disservice if test items from relevant cognitive levels are not included in the name of difficulty. That becomes only an excuse and disservice. At the same cognitive level, we can tone down the difficulty of the assessment item.

(Refer Slide Time: 36:41)



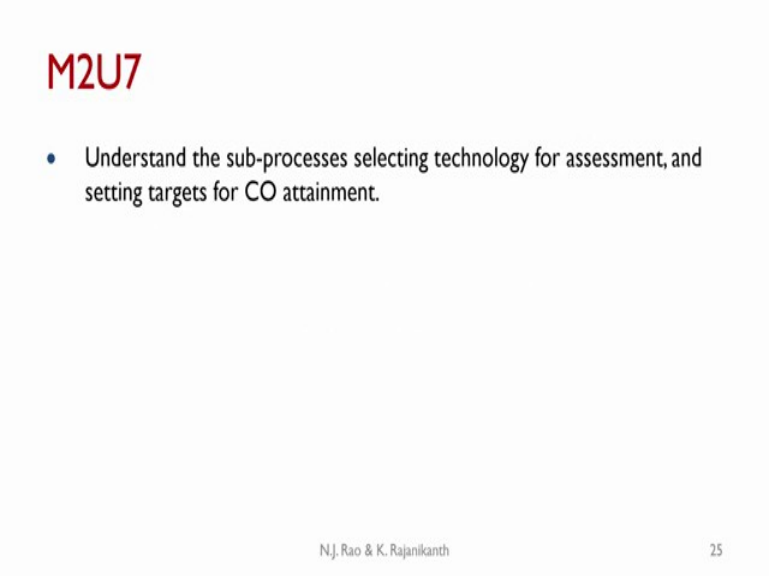
The sub-processes of a design phase in the context of an engineering course: The proposed sub-processes and their sequence are - first selecting the technology for assessment and evaluation, (which we will see in the next unit); Setting targets for CO attainment, designing the assessment pattern and assessment instruments. This must be done upfront to ensure quality of assessment, which ensures quality of learning; creating the item bank.

(Refer Slide Time: 37:13)



Exercise: list aspects of assessment not addressed in this unit, but still considered relevant to your course. Describe any different sub-processes you consider necessary to be included in the Design Phase of ADDIE with respect to designing your course. Thank you for sharing the results of this exercise at tale.iista@gmail.com.

(Refer Slide Time: 37:41)



In the next unit we will discuss the design phase sub-processes. Outcome of that unit would be: understand the sub-processes of selecting technology for assessment and setting targets for CO attainments. This selecting the target for CO attainment is again a crucial aspect of the NBA process of quality improvement. We set the target for the CO attainment up front, we deliver the course - design the instructional delivery to attain those levels; we measure the actual attainment levels through assessment. If the actual attainment level is as per the target, then we can think of increasing the attainment level targets, but if you fail to meet the intended target levels, then we have to look into the reasons for that and try to create remedial measures, improvement plans so that the next batch does achieve the intended target levels - that will be completing the quality loop.

This is the essence of improving the quality of learning by the students. To do this it is very essential that the assessment must give us data which is really useful for us for closing this quality loop, which essentially means that the assessment must be valid and reliable. That requires certain process to be followed, that is what we have been discussing in the design phase and we will continue with this in the next unit.

Thank you very much and we will meet you again.