Introduction To Multimedia Applications
CT801-4-0-OIMA

ASIA PACIFIC UNIVERSITY
OF TECHNOLOGY & INNOVATION

Evaluation of Multimedia Applications

Topic & Structure of Lesson

To discuss the need for evaluating multimedia applications.

To discuss the 2 approached used to evaluate multimedia application.

To explain the evaluation framework proposed to evaluate a multimedia application.

Learning Outcomes

At the end of this lecture, you will be able to:

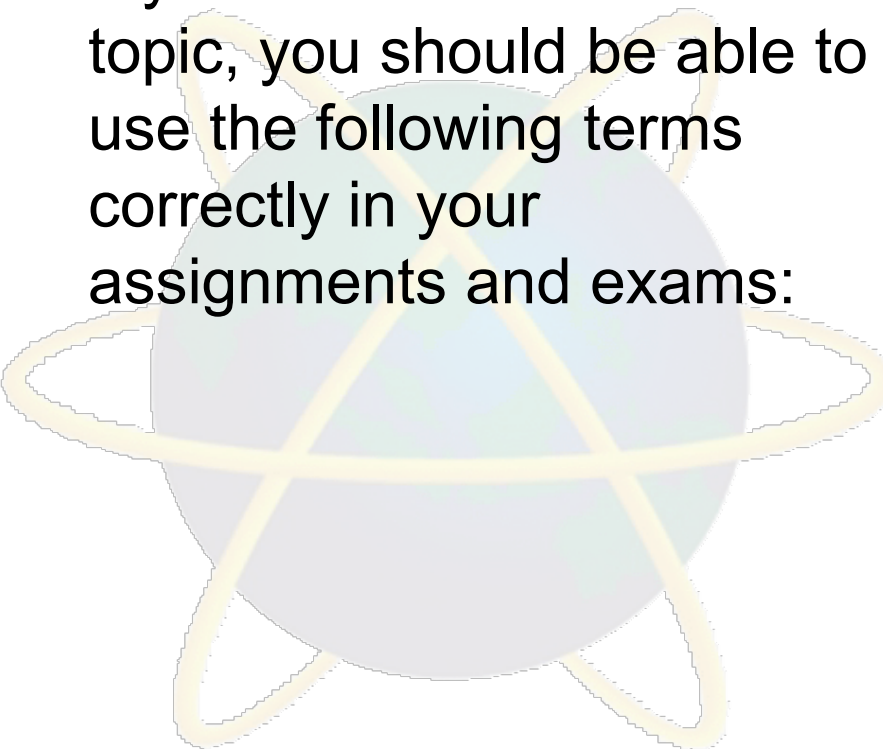Discuss, explain and differentiate summative evaluation and formative evaluation

Outline and use an evaluation framework for evaluating multimedia systems

Discuss the processes involved in each of the stage in the evaluation framework

Key Terms you must be able to use

If you have mastered this topic, you should be able to use the following terms correctly in your assignments and exams:

Summative evaluation

Formative evaluation

Evaluation framework

Evaluation Tools (Quantitative & Qualitative)

The definition of "Usability"

Usability of Multimedia applications :

Does the user feel in control of the application ?

To what extent can the user achieve their goals using the application?

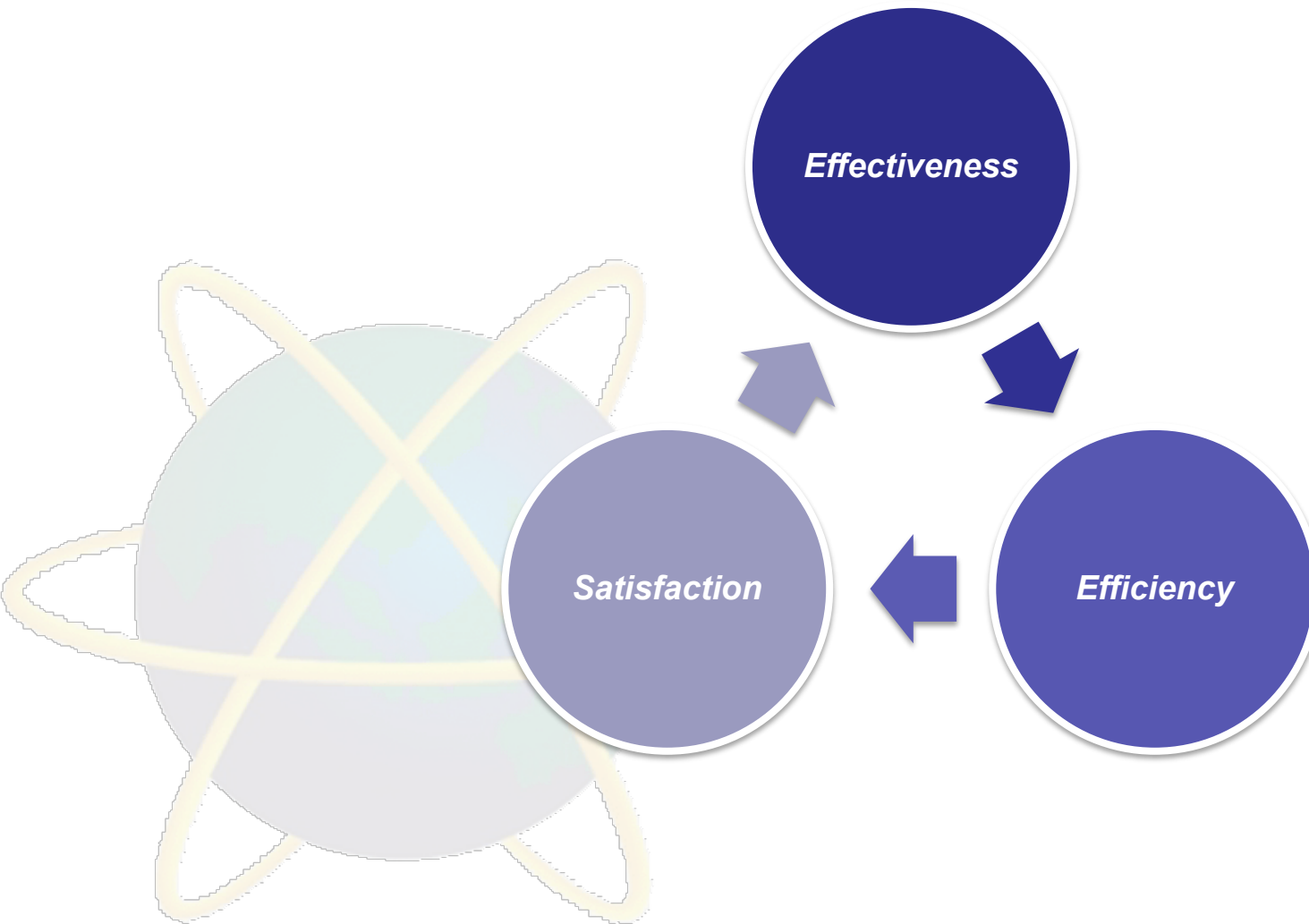How far does the product appear to assist the user ?

How easy is the application to learn ?

How does the user respond emotionally ?

The important 'Dimensions' of Usability are therefore :

✓ *Effectiveness , Efficiency, Satisfaction*

Dimensions of Usability

Formative evaluation

Formative

- **Particularly suitable for evaluation on real-work situations (real time-basis/at that particular moment)**
- Usually qualitative will be used for feasible study and post-implementation requirements
- Suitable for new program/activity or early implementation

Contextual inquiry

Co-operative/Peer observation

Summative evaluation

Summative

- **Focusing on the final outcome/result**
- Produces quantitative and qualitative data about the finished system.
- Suitable for instructional design and overall performance & accountability

Usability & quantitative metrics
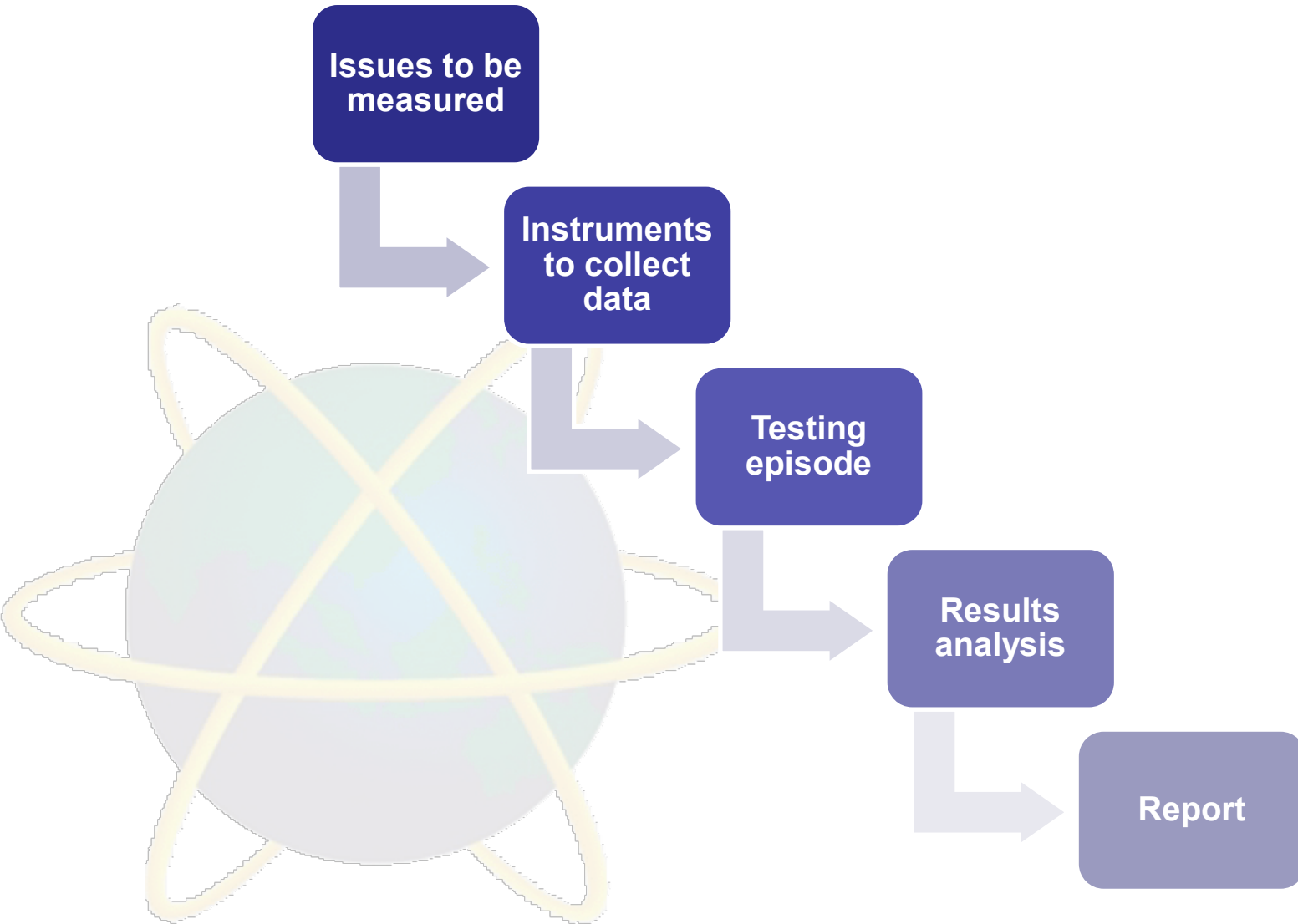
Psychometric & psychology tests

In-depth interviews & pattern analysis

Evaluation Overview

Evaluation is required to find out how well a multimedia interface works for a user;

To do this an evaluation framework is required

Evaluation Steps

Issues to be measured → **Instruments to collect data** → **Testing episode** → **Results analysis** → **Report**

Evaluation:
1. Issues Measured

Measurements can be user related, e.g. -

- Attitude, user satisfaction (positive or negative)
- Knowledge, recall (what has been learned)
- Goal related (aim to achieve a task, effectiveness)
- Usability criteria based (is the system usable)
- Learnability
- Performance based (efficiency)
- Error rates

Measurements can be system related, e.g. -

- Software structure
- Capture paths taken through the system and features used
- Response, or time delays
- Highlight navigation problems (lost in hyperspace)

Evaluation:
2. Instruments

The data collection technique used in the testing episode.

Example aproaches that involve end users are -

- i) Semi-structured interviews
- ii) Questionnaires
- iii) Incident diary, self-reporting
- iv) Feature checklist
- v) Focus groups
- vi) Think aloud protocol
- vii) Experiments
- viii) Usability Laboratories

Evaluation: Instruments

Semi-structured interview	**A quantitative / retrospective method (depending on design)**	**MUMMS - Measuring the Usability of "Multi-Media" Systems**	**Incident diary or 'self-reporting' method**

• A qualitative / retrospective method
• Uses an agenda of questions
• Can focus on specific issues of interest
• Time intensive, only suitable for small numbers
• Good for discussing interface options, what's good / bad, and suggesting improvements
• Could also use an interview to 'pilot' a questionnaire
	• Structured questions, answers typically (not exclusively) in the form of
	  • agree/disagree/neutral yes/no 1/2/3/4/5
	  • 'Rating' scale
	• Requires little time to administer once designed, user can fill in on their own, suitable for large numbers
		• Contains a number of 'subscales' for measuring end-user perceived quality of systems, including the extent to which the user feels they are in control of the 'pace'.
		• A new scale currently called 'excitement' is being considered - the extent to which users feel 'drawn into' the application's world, (fascination).
			• Quantitative/on-the-spot
			• A structured diary for logging incidents
			• To catch interface problems that have been missed by other instruments, or that cannot be simulated in a lab environment,
			• As an alternative to think aloud protocols if this is uneconomical (too many subjects) or impractical (investigator can't be there).
			• Good for: finding problems with interface, time delays, occurrence of being lost in hyperspace

Evaluation: Instruments

Feature checklist	Focus groups	Think aloud protocol	Experiments (controlled experiments)	Usability laboratory
• Quantitative / retrospective • Examines features used • Usage, knowledge required, need • Takes 2-15 minutes • Good for : what hypermedia facilities are used, and node access	• Qualitative / retrospective method • Organised as a group discussion • Works on the concept of human triggers (someone says something, others pick up) • Time intensive normally about an hour • Good for: interface options, what's good / bad, and suggesting improvements	• Qualitative / on-the-spot method • Use software, and record spoken user views as they use the application to perform a task • Can reveal how, not just where, people get stuck • - Normally requires about an hour (any longer & there's loss of concentration etc.) • Good for finding out how a system is used, and problems related to the system	• Quantitative / on-the-spot • Good for gathering lots of different types of information • Gaining empirical evidence to support a claim or hypothesis • Typically lasts 1 to 3 hours (However, may need to run an experiment several times) • To begin an experiment we need to create a design plan by first forming a hypothesis; the idea, combined with a null hypothesis, Choose 'Subjects', Select 'Variables' and pilot the experimental design	• 'Office' environment with 4-5 desks, computers etc. in soundproof room ('Test Room'), video cameras etc. • 'Observation room' with 1-way mirror • 'Control room'

Evaluation: 3. Testing episode

Design the test

Use the instrument (s) to obtain results

Support for results capture

- **Paper & pencil (limited and hard work)**
- **Audio recording (improvement slightly limited)**
- **Video recording (advanced)**
- **User note books (good for unsupervised user)**

Computer logging

- **- Excellent for multimedia applications**
- **- Node access, errors (and where made), time spent at nodes etc.**

Evaluation:
4. Results

Analyse the results obtained

Quantitative - what's going on, statistical analysis

- **E.g. Instruments**
- **Experiments, questionnaires, feature checklists, incident diaries**

Qualitative - explain what's happening

- **E.g. Instruments**
- **Think aloud, focus group, semi-structured interview**

Quick Review Question

01
Give some examples of where formative assessments can be applied
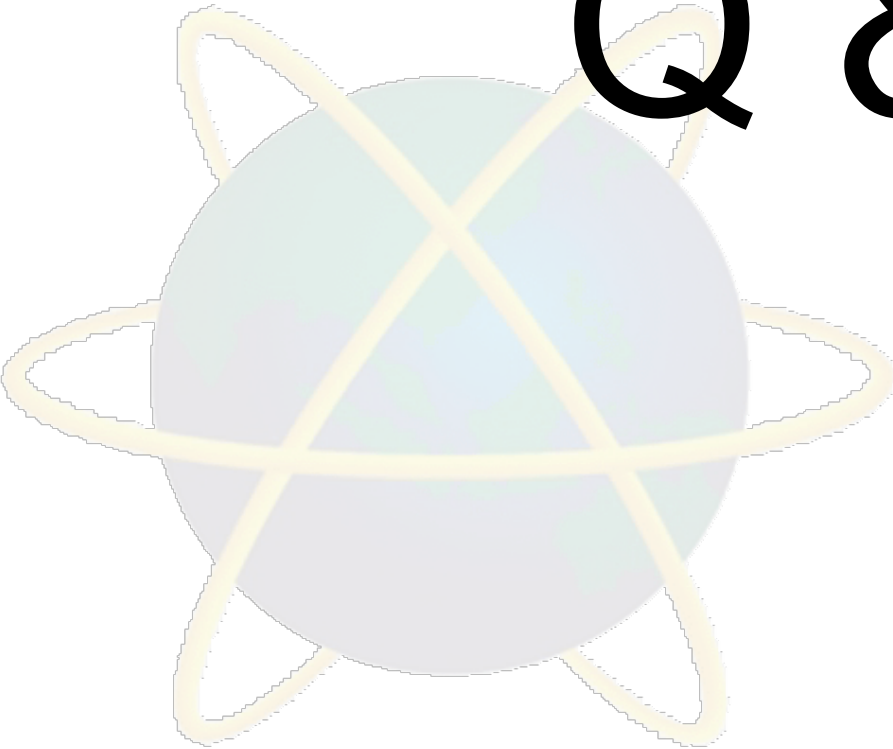
02
Give some examples of where summative assessments can be applied.

03
Explain differences between the two assessments.

Question and answer session

Q & A

Next Session

Authoring Tools