Evaluation and Testing (IO5)

Louise Fryer

Utopian Voices Ltd.

Intellectual outputs

- **O1**: Gathering EU best practices (UAM)
- **O2:** Defining audio-describer competences (UNITS)
- **O3**: Producing course design (UA)
- **O4**: Creating training materials (UAB)
- **O5**: **Testing and evaluating training materials (UV)**
- **O6**: Attributing ECTS/ECVETS to materials (UNITS)

- **Full title**: Evaluation and Testing
- **Duration** initially envisaged running intermittently M7 – M36
  - In fact: M1 – M36 (Oct 16 – Sept 19)
- **Leading partner**: UV (Fryer)
- **Contributing partners**: all
- **Other contributors**
  - All participants in ADLAB PRO events (AD providers, users etc.)
  - AD students, potentially trainers/lecturers/course deliverers

Evaluation for ADLAB PRO

- Evaluation ÷ Internal (Management/Project evaluation)
  - External (Output evaluation)

According to the application form

Evaluation should be built in from the start

Course content (IO4) will be designed in progressive form (tested at each stage to secure content quality, adequacy and progress.)

Assesment of Outputs 1 – 4 (short reports) (so far IO1 & 2 reports completed)

Partners to discuss

- ADLAB PRO evaluation cycle
- IO4 testing & evaluation

What is Evaluation?

- "a data reduction process that involves the collection of large amounts of data which are analysed and synthesized into an overall judgement of worth or merit." Wigley (1988, p.21)

Why evaluate?

"Evaluation is given a low priority in the instructional process, a contention that is supported by the small number of articles in the literature that deal with it."

Evaluation should be given a high priority as it provides evidence that justifies the value and viability of training programmes.

"evaluation is often something of an afterthought for those whose main concern is with delivering training" (Marsden,1991, p.31 cited in Foxon, 1989).

AIMS

- To improve a project.To check that it's meeting its goals
- To better communicate its achievements
  - Serendipitous gains/unexpected insights
  - To feed into interim reports
  - To show what we've done & to celebrate what we've achieved

There are 2 types of evaluation/Assessment

- **Formative** (assessing ongoing activities) (implementation and progress)

- **Summative** (assessing the end result)

- Formative: the chef tasting the soup as (s)he makes it
- Summative: everyone else tasting the soup when it's ready

Formative assessment in ADLAB PRO

- **Design, distribution, completion & analysis** of questionnaires for
- RM evaluation
- TPM evaluation
- ME evaluation
- IO evaluation

I like to think of assessment as...

or

For example

- a comment on the EF after RM3 suggested one partner was unhappy and thinking of leaving.
- Swift action from lead partner (UNITS) allayed the concerns & the partner stayed.

Summative assessments in ADLAB PRO so far

- RM reports (12)
- TPM reports (2)
- ME reports (2)
- IO reports (2)

Advantages of Formative Evaluation

- Everyone's voice is heard, equally and often

- provides easy to access data

- combines quantitative and qualitative data

Limitations

Formative evaluation can be seen as:

- Unnecessarily bureaucratic
- Time-consuming
- Repetitive➜fatigue
- Lack of anonymity
- Small sample size
- Ceiling effects

Steps taken to minimise the limitations

- Work hand in hand with the Quality Manager to avoid repetition
- Tick box format for most indicators, single page format
- Mixed measures helps overcome the ceiling limitations
- Reports kept short
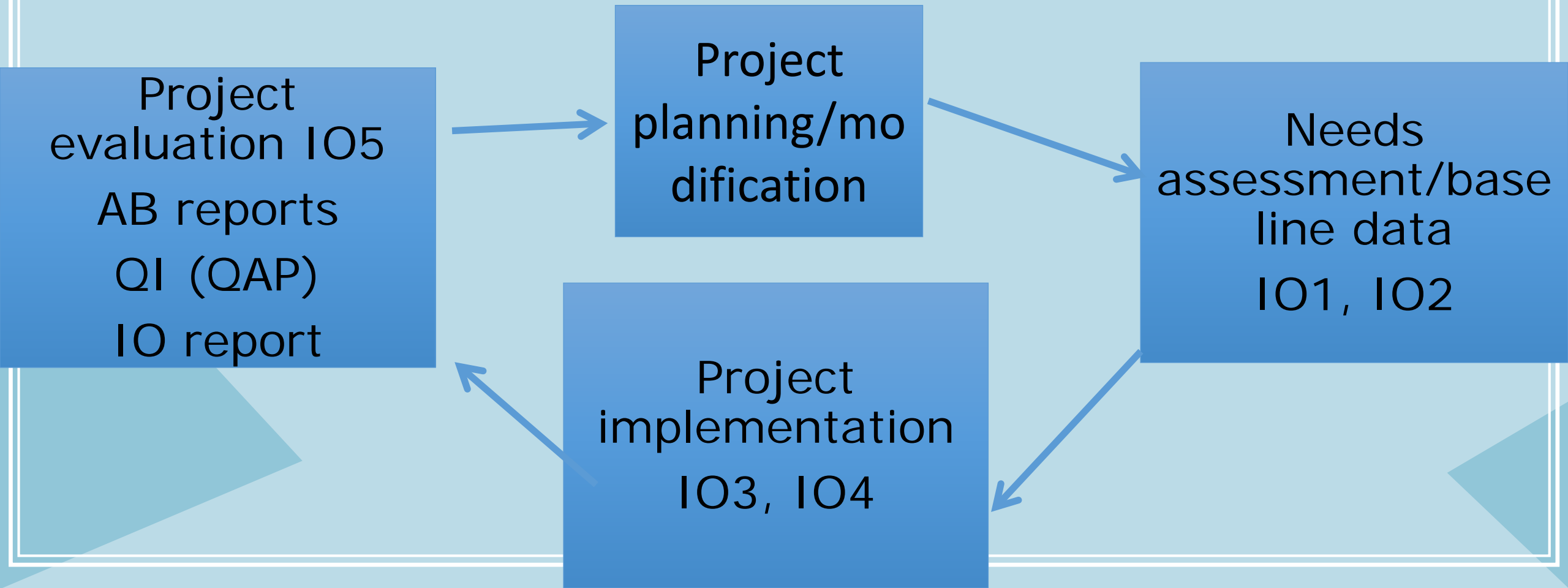- by synthesising data, formative evaluation should actually save time

Planning /Evaluation Cycle
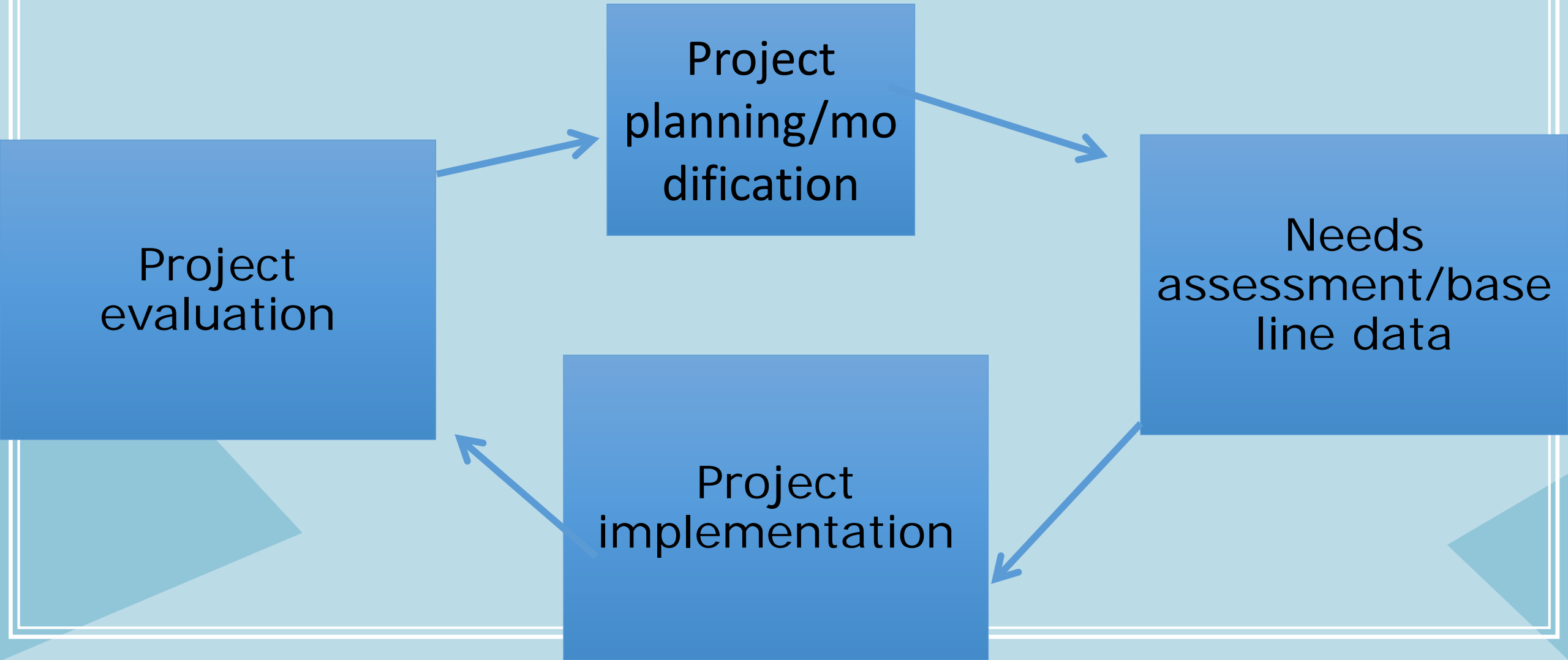
IO testing and evaluation

- Quality Report for each IO comprising:

- Summary of AB evaluations (Praise; criticism; recommendations)

- **Quantitative Quality Indicators (taken from the QAP)**:EF completed by Partner responsible for evaluation/validation.

- Leading to a mark /60

- List of publications/ presentations relating to that IO

The big question

?
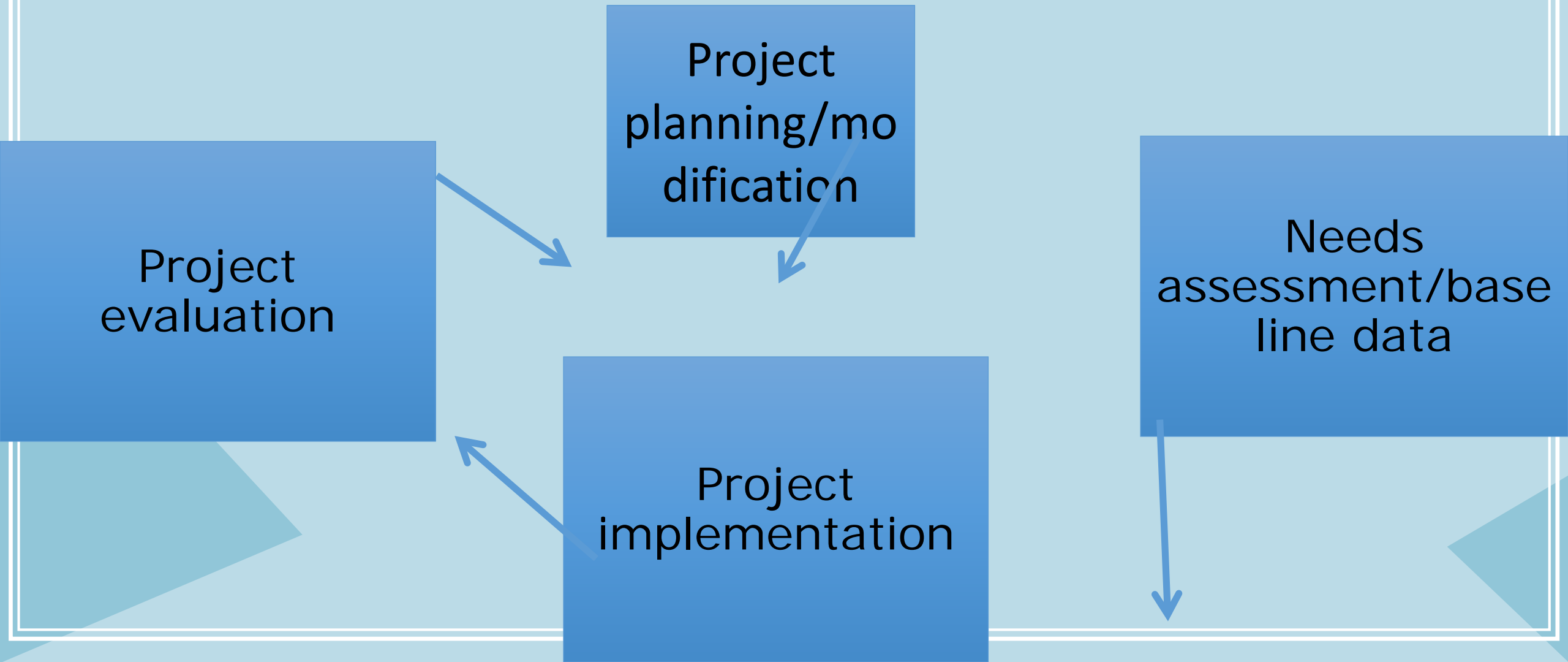
What happens to the evaluation reports?

- RM/ME/IO evaluation forms➜ data analysed➜evaluation report➜?

- AB IO evaluation reports➜?

Is our Planning /Evaluation Cycle broken?

Project planning/modification

Project evaluation

Needs assessment/baseline data

Project implementation

What we're good at

- Capturing data

What we could improve

- Acting on the data

e.g. recommendations from AB for IO2 with potential to modify IO2 & affect IO's 3 & 4

- I felt the need for some summary and highlighting from time to time. (BB)

- It would be interesting to explore the extent to which AD training can be considered "generic", i.e., its fundamentals apply to all formats in which it is practiced (performing arts/museums/media).

- a questionnaire – in my view – should not be the only way to build the project/course on. Therefore, I believe the project should address the issues of learning, acquiring skills and developing competence, focussing on the specificities of AD.

Recommendations ctd.

- What are the prerequisites (abilities?) necessary to become an audio describer? For instance, in the case of the translator, it would be fluency in two languages. It would be helpful to see how these **abilities** develop into **competences** and then into **professional expertise**. Is there a progression path envisaged in the AD training course? (AS)

- The report does mention the social-constructive teaching model and project-based learning (Kiraly 2000), which are very important in contemporary translator (and – by extension – audio describer) training, but the question of audio describer competences (skills?) is only mentioned in passing. Given the high-profile of the project and a great intellectual potential of its research team, I would suggest developing a competence framework, following the example of the EMT or PACTE models in translation studies.

How are we going to improve?

- Suggestions:

1: UV alerts lead partner for the next IO when the Report for the previous IO is ready.

- The lead partner needs to demonstrate how recommendations have been carried forward in the new IO.

2: Build in time for revision in response to AB reports.

furthermore

- A proposal: No new IO/ME/meeting is organised before the previous report has been read & recommendations acted upon. Or justifications given as to why the recommendations should not or cannot be acted upon.

What's gone well

- Modifying ME's in response to feedback

Still to come: collaboration with UAB for IO4

- **"testing will take place from the very beginning of the cycle**, so that each new resource added will benefit from measures of user knowledge and feedback, gained from testing of existing ones."

-

What does that mean in practice?

- The tasks, proposed in IO3, need to be assessed by users (students?)before they are adopted by partners.

- They will comprise: a baseline measure to be completed before and after the task (what the student already knows vs what they have learned to demonstrate "value added")

PROPOSED measures

- The task has improved my understanding of the needs of AD users
- The task was well structured
- I found the task interesting
- I found the task confusing

The task has increased my understanding of audio description

The demands of the task were appropriate

where 1= minimal effort and 9 = extreme effort, how much mental effort did it take to complete the task (a simple measure of cognitive load, Paas,1992)

IO4 evaluation

- "Users" should assess the tasks but does that mean academic students? Vocational students? Trainers? Lecturers? Current course providers? Potential course providers?

- Will users assess all tasks?

- What constitutes a task? Is a ppt. a task?

- We need to know how long the "tasks" take (for IO6) but do we control this in any way? Or can participants take as long as they want?

Evaluating student progress for pedagogical purposes.

UV has produced and been testing EFs for assessing student performance on 4 measures of Quality (delivery; accuracy; language; synchrony). A chapter explaining the rationale is in press: Quality Assessment in Audio description: Lessons learned from Interpreting. In E. Huertas-Barros, S. Vandepitte and E. Iglesias-Fernández (Eds.) Quality Assurance and Assessment Practices in Translation and Interpreting. IGI-Global

What partners could do for evaluation in IO3 & IO4

- Share their current evaluation practice (academic partners)
- Maybe AMU could ask some contributors to IO1 (trainers) to assess individual modules in IO3 once they're ready?

Today we should decide...

- What constitutes a "user" in IO4
- What constitutes a "task"?
- How do we determine duration?
- How can non-academic partners contribute to assessment and evaluation of IOs 3 & 4?

We also promised...

- Other forms of **qualitative feedback**, such as **focus groups and one-to-one interviews** with service users will take the form of quality assessment of audio descriptions prepared by course participants at the end of the course.

- Are we going to test the whole course?

Points for UV

- Make forms & reports easier to find (Laura has been helping. Thanks!)
- More consistent naming
- Alert partners when reports are ready
- A greater emphasis on collecting baseline data
- A greater emphasis on qualitative data from MEs etc.
- Others?

Final output of IO5

- short reports ultimately combined into a handbook (print-based and online), satisfaction survey forms and measurable satisfaction indexes, **methodological guidelines** for testing cognitive load and other reception variables

- I've made a start on this as a "how to" guide for future projects.

references

- Blank,R.(193) Developing a System of Education Indicators: Selecting, Implementing,and Reporting Indicators. Educational Evaluation and PolicyAnalysis,15(1,Spring):65-80

- Foxon, M (1989) Evaluation of Training and Development Programs: A review of the literature. Australasian Journal of Educational Technology, 5 (2)

- Frechtling, J. (2002) The 2002 User-Friendly Handbook for Project Evaluation