## Language & IT Dr. Abdullah Al Fraidan

## **Judgmental Evaluation**

## Lecture-4-

## التقييم الحكميThe judgmental evaluation

- Methods of evaluation (A): Introspective judgmental evaluation; checklists
  - أساليب التقييم (A): تقييم حكمى الاستقراء، وقائمة المراجعة
- There are two broad types of way of actually executing evaluation studies (A and B here).
  - هناك نوعين واسعين لوسيلة تنفيذ دراسات تقييم الواقع (A و B هنا).
- In many ways A suits situations 1 and 2 above, B suits situation 3. (Cf Chapelle 2001 p53).
  - في نواح عديدة -A- تناسب الحالات 1 و 2 أعلاه، B تناسب الوضع 3. (راجع شابيل 2001).
     P53
- Introspection means relying on one's own judgment/experience, and maybe published consensus on what should be there, what is good or bad, or AL theory.
- Introspection تعني الاعتماد على الحكم بنفسه / الخبرة، وتوافق الآراء المنشورة ربما على ما ينبغي أن يكون هناك، ما هو جيد أو سيئ، أو AL النظرية.
- (A1) Evaluation can be done purely individually, subjectively, globally and introspectively. I.e. the teacher simply looks through the material, or in our case tries out the program (or just reads the blurb about it in a catalogue), and comes to an overall intuitive judgment about whether it would suit their class or what class it would suit.
- (A1) يمكن التقييم أن يتم بشكل فردي بحت، ذاتيا، عالميا. أي يبدو المعلم ببساطة من خلال المواد، أو في حالتنا يحاول الخروج من البرنامج (أو يقرأ مجرد دعاية مغالى فيها عن ذلك في نشرة مصورة)، ويأتي إلى الحكم بديهيا العام حول ما إذا كان من شأنها أن تناسب فئة أو فئة ما من شأنها أن تناسب

- When teachers evaluate in this way it may help in part to try to place themselves in the role of some type of learner using the material.
  - عندما تقييم المعلمين بهذه الطريقة التي قد تساعد جزئيا إلى محاولة لوضع أنفسهم لدور بعض نوع من المتعلم باستخدام المواد.
- When trying out a CALL program it is especially useful often to make deliberate mistakes to see how the program responds - e.g. give wrong answers and press the wrong keys etc.
- عندما تحاول الخروج بالبرنامج -غالبا من المفيد بشكل خاص إلى ارتكاب أخطاء متعمدة لنرى كيف يستجيب البرنامج على سبيل المثال إعطاء إجابات خاطئة والضغط على مفتاحي خاطئ الخ
- This could be described as the global 'expert judgment' method of evaluation.
  - يمكن وصف ذلك بأنه أسلوب عالمي "حكم الخبراء" من التقييم.
- The evaluator introspects and somehow accesses an unanalyzed notion of some users of the software, an unanalyzed impression of the software, and matches the two using often inexplicit criteria.
  - المقيم بطريقة أو بأخرى يصل إلى فكرة لبعض مستخدمي البرنامج، والأنطباع من البرنامج، يتطابق مع اثنين باستخدام معايير غالبا ضمنية.
- (A2) However, to regard evaluation as in any way systematic it is necessary at the very least to 'unpack' this armchair approach a bit.
  - مع ذلك، على اعتبار التقييم كما الحال بأي طريقة منهجية فمن الضروري على الأقل إلى 'فك' هذا النهج المقاعد قليلا.
- The teacher (or anyone else) acting alone as evaluator should break down the 'overall' or global judgment into parts.
- يجب على المعلم (أو أي شخص آخر) يعمل بمفرده كما المقيم لكسر الحكم 'العام' أو عالمية إلى أجزاء.
- This means (a) looking carefully at different aspects of the materials separately and (b) thinking of all the relevant different aspects of the learning situation, learners, potential use etc. etc. and (c) judging aspects of (a) in respect of (b), broken down into points.
- هذا يعني (A) تبحث بعناية في جوانب مختلفة لمواد على حدة و (b) للتفكير من جميع الجوانب (a) ذات صلة مختلفة لوضع التعلم، والمتعلمين، واستخدام إمكانات الخ و (c) للحكم بجوانب (a) فيما يتعلق من (b)، مقسمة إلى نقاط

- This last in part resembles the process of assessing 'content validity',
   often talked about in language testing: one can check on an achievement
   test by analysing the aspects of language tested and comparing them
   with what the syllabus or the teaching course before the test covered.
  - هذا الأخير في جزء يشبه عملية التقييم "صلاحية المحتوى"، وغالبا ما تحدث في اختبار اللغة: يمكن للمرء أن يتحقق على اختبار التحصيل من خلال تحليل جوانب اللغة لأختبار ها ومقارنتها مع المنهج أو أثناء التدريس قبل تغطيت الاختبار
- Another general principle of language testing also applies here: it is known that tests with more items are more reliable than shorter ones, and a set of agree/disagree items circling round some issue is more reliable than a single one targeting it.
- أيضا ينطبق مبدأ عام آخر لأختبار اللغة هنا: أنه من المعروف أن الأختبارات مع أكثر البنود هي أكثر موثوقية من الأقصر منها، ومجموعة من البنود موافق / غير موافق تدور حول بعض المسألة لأكثر موثوقية من أي واحد تستهدفه.
- So here, the summary of a whole series of introspective judgments of specific aspects is more reliable than one global one.
  - حتى هنا، ملخص سلسلة كاملة من أحكام الاستقراء لجوانب محددة هو أكثر موثوقية من واحد عالمي.
- This is where 'checklists' come in.

- هذا هو 'قوائم' تأتي حيث فيها.
- These are written records of the sort of 'breakdowns' just described.
  - تتم كتابة هذه السجلات لهذا النوع من 'انهيار' وصفه للتو.
- They may be made by the teacher/evaluator, or adopted from someone else.
  - قد تكون مصنوعة من قبل المعلم / المقيم، أو اعتمد من شخص آخر.
- They at least provide a way of ensuring that important aspects do not get forgotten and that there is some consistency if the same person evaluates several things.
  - أنها على الأقل توفر وسيلة لضمان أن الجوانب الهامة لا تحصل على حالنا وأنه هناك بعض الاتساق إذا كان الشخص نفسه يقييم العديد من الأشياء.
- However, the evaluation still remains individual, introspective and maybe pretty subjective.
  - مع ذلك، فإن التقييم لا يزال فردي، أستقرائي، وربما شخصي جدا.

- Checklists generally take the form of sets of headings to be considered or sets of questions to ask oneself.
  - عادة ما تتخذ قوائم المراجعة شكل مجموعات من العناوين التي يتعين النظر فيها أو مجموعات من الأسئلة للتساءل .
- They may or may not include a system for weighting different elements, or adding up a total score in some way.
  - أنها قد أو قد لا تتضمن نظاما لترجيح العناصر المختلفة، أو إضافة ما يصل إلى النتيجة الإجمالية في أو بأخرى.
- Two I know of for CALL are the list of points in Jones and Fortescue, and a more reasoned and systematic framework by Odell (in Leech and Candlin).
  - أنا أعرف اثنين من ال CALL هي قائمة من النقاط في جونز وفور تسكو، ووضع إطار أكثر سبب ومنهجي من قبل أوديل.
- Recently Chapelle has a set of 6 points formed from an SLA research perspective (2001 p54ff).
  - مؤخرا شابيل لديه مجموعة من 6 نقاط شكلت لمنظور البحث . SLA
- John Roberts has a much bigger collection of such checklist used in general materials evaluation.
  - جون روبرتس لديه مجموعة أكبر بكثير من تلك المرجعية المستخدمة في تقييم المواد العامة.
- However, many published checklists strike one as a rather miscellaneous collection of points or questions, not clearly distinguishing between (a) and (b) and (c) above, and not obviously exhausting the types of point that should be considered, or organising them in a motivated way.
  - مع ذلك، العديد من قوائم المراجعة التي نشرت ضربت كواحد لمجموعة متنوعة بدلا من نقاط أو أسئلة، لا تميز بوضوح بين (a) و (b) و (c) أعلاه، وليس من الواضح استنفاد أنواع من النقطة التي ينبغي النظر فيها، أو تنظيمها بطريقة دوافع.
- For teachers, often the checklist-based evaluation just described is the only one feasible, since it is the one that can be done quickly and easily and before the materials have been extensively used or even bought.
- غالبا ما يكون تقييم المعلمين على أساس قائمة صفتها ليست سوى واحدة ممكنه، لأنه هو الذي يمكن القيام به بسرعة وسهولة وقبل المواد التي استخدمت على نطاق واسع أو حتى تشترى

- It can be enhanced by incorporating the views, arrived at in a similar way perhaps, of more than one person. I.e. the teacher can get other teachers to do the same sort of evaluation, or read reviews in journals etc.
  - يمكن أن تتعزز من خلال دمج وجهات النظر، ربما وصلت بطريقة مشابهة ، من أكثر من شخص واحد. أي يمكن للمعلم الحصول على معلم آخر للقيام بنفس النوع من التقييم، أو قراءة أراء في المجلات الخ
- This makes it less individual, though still introspective and rather subjective.
  - هذا يجعلها أقل فردية، وإن كان لا يزال أستقراء وذاتية لحد ما.
- (A3) Additionally the teacher may enhance the checklist approach, if he/she has the time and energy....., by doing things that in a loose sense could be called 'research'.
  - (A3) بالإضافة إلى ذلك المعلم قد يعزز نهج المرجعية، واذا كان / كانت لديه الوقت والطاقة .....، عن طريق القيام بأشياء أنه بمعنى سائب يمكن أن يسمى 'بحث'.
- By this I mean looking systematically with some analytic techniques etc. at aspects under the (a) or (b) head above, not just deciding what they are on an instant introspective basis.
  - أعني بذلك تبحث بشكل منهجي مع بعض التقنيات التحليلية الخ في الجوانب تحت (a) أو (b) أعلاه ، وليس فقط لأتخاذ قرار ما هي عليه على أساس الاستقراء الفوري.
- This may focus more on the (a) side: e.g. linguistic analysis of the structures used in the content of the program (if it is fixed), checking the frequency level of the vocabulary against a standard reference list, grading the exercise types that are incorporated on a recognised scale of task difficulty etc.
  - قد تركز أكثر على هذا الجانب (a): على سبيل المثال التحليل اللغوي من الهياكل المستخدمة في محتوى البرنامج (إذا هي ثابتة)، والتحقق من مستوى التردد من المفردات ضد قائمة مرجعية معيارية، وتدرج أنواع التمارين التي يتم إدراجها على نطاق ومعترف بها لصعوبة المهمة الخ

- This might be called 'materials analysis'. Or it may focus on the (b) side:
   e.g. finding out what the syllabus for the current year actually says my
   learners should be doing, doing an analysis of learners' needs or
   interests, finding out what the school budget actually has available, etc.
  - هذا يمكن أن يطلق عليه "تحليل المواد '. أو أنها قد تركز على الجانب (b): على سبيل المثال معرفة المنهج للعام الحالي يقول في الواقع المتعلمي يجب أن يقوم به، والقيام بتحليل احتياجات أو مصالح المتعلمين، ومعرفة ما لديه ميزانية متاحة فعلا للمدرسة ، الخ
- This is in effect 'analysis of the learning/teaching situation'. These are all things that might appear on a checklist and of course can all alternatively be decided by the evaluator just "off the top of his/her head".
- هذا هو تأثير في "تحليل التعلم / وضع التدريس ". هذه هي كل الأشياء التي قد تظهر على قائمة المرجعية وبالطبع يمكن للجميع وبدلا من أن يقررها المقيم فقط "من أعلى / رأسها".
- Further, with respect esp. to (c) the suitability judgment itself, these may bear some 'research' in the form of reading up what theory, research studies and so forth have to say.
  - علاوة على ذلك، مع تقدير ESP. إلى (c) بالحكم لملاءمة نفسه، وهذه قد تحمل بعض 'البحث' بشكل القراءة بما للنظرية، وهكذا يجب أن أقول الدر اسات البحثية .
- You have a program with certain characteristics and you want to use it with young learners (as the publishers indeed claim it is suited to be).
  - لديك برنامج لخصائص معينة، وترغب في استخدامه مع المتعلمين الصغار (كما في الواقع يدعون الناشرين مناسبة لها أن تكون).
- Instead of just relying on one's own judgment of what is suitable, one can read up what the collective wisdom of psychologists, educators etc. have to say about what the characteristics are of young learners and so what suits them.
- بدلا من الاعتماد فقط على حكم المرء لما هو مناسب، يمكن للمرء أن يقرأ ما الحكمة الجماعية من علماء النفس والمربين الخ يجب أن أقول حول ما هي خصائص المتعلمين الصغار وحتى ما يناسبهم.
- Similarly the general wisdom on how to construct multiple choice items (e.g. in books on testing) may help evaluate the suitability of m/c items in a CALL package.
  - بالمثل، فإن الحكمة عامة على كيفية بناء وحدات متعددة الخيارات (مثلا في الكتب بشأن الاختبار) قد تساعد على تقييم مدى ملاءمة البنود c/m لمجموعة CALL.

- Research studies of the way learners use CALL, teaching with CALL etc. may also be worth looking at, and indeed if a program is supposedly designed to aid reading, the general wisdom on the teaching of reading and reading strategies, and so forth.
  - الدراسات البحثية لطرق المتعلمين بأستخدام CALL، والتدريس مع CALL وما إلى ذلك قد يكون أيضا يستحق النظر، بل ويفترض إذا تم تصميم البرنامج لمساعدة القراءة، والحكمة العامة على تدريس القراءة واستراتيجيات القراءة، و هكذا
- However, there is always the danger that supposedly 'general' research findings do not actually apply in your situation for some reason.
  - مع ذلك، هناك دائما خطر أن نتائج البحوث من المفترض أن 'العام' لا ينطبق في الواقع في وضعك لسبب ما.
- But if you are using the checklist approach there are some key things not to forget:
  - ولكن إذا كنت تستخدم نهج المرجعية لا ننسى ان هناك بعض الامور المهمة:
- Be explicit about where the list comes from, which existing one is being used/adapted, and have as many detailed subsections as possible.
  - كن واضحا حول ما يأتي من القائمة، والذي يتم استخدامه موجود / تتكييف، ولها العديد من الأقسام الفرعية مفصلة بقدر الإمكان
- Make sure whatever system/list you use covers all three of the

   (a) (b) and (c) aspects Cover the (a) aspect. A description of
   detailed aspects of how the program works, with examples of
   actual items, screens etc., and what it does (a) has to be
   incorporated, since the reader cannot be assumed to be familiar
   with the software.
- تأكد من أي نظام / قائمة تستخدم لتغطي كل ثلاثة من (a) و (b) و (c) كجوانب تغطية للجانب (a). وصف الجوانب التفصيلية لكيفية عمل البرنامج، مع أمثلة من البنود الفعلية، والشاشات وما إلى ذلك، وماذا يفعل(a) ، لأن القارئ لا يمكن أن يفترض أن يكون على دراية مع البرنامج.

- If part of what you are evaluating is a particular task that is not part of the software itself, or some language element supplied by the teacher, make that clear.
  - إذا كان جزء ما يقييم مهمة خاصة ليس كجزء من البرنامج نفسه، أو بعض عنصر اللغة الموردة من قبل المعلم، ويجعل ذلك واضحا.
- But that alone is not an evaluation.

- لكن ذلك وحده ليس تقييما
- Cover the (b) aspect. Give a full account of (imagined or real) target learners in a situation in a particular country at a particular level etc.
  - تغطية الجانب (b). إعطاء الاعتبار الكامل (المتخيلة أو الحقيقية) هدف المتعلمين بوضع في بلد معين عند مستوى معين الخ
- Evaluation for some generalised 'learner' is not very convincing.
  - التقييم لبعض تعميم "المتعلم" ليست مقنعة جدا
- Don't forget (c) i.e. explanation of how each feature of the program (a) does or doesn't fit (b).
  - لا ننسى (c) أي شرح لكيفية كل ميزة من البرنامج (a) أو لا تتناسب مع (b).
- This needs to be supported wherever possible by more than your expert intuition - reference to applied linguistic concepts, research, models etc. (E.g. Chapelle 2001 pp45-51).
  - هذا يحتاج إلى أن يدعم حيثما كان ذلك ممكنا من قبل أكثر من حدس لخبير خاص إشارة إلى المفاهيم اللغوية التطبيقية، والبحوث، والنماذج الخ .
- This is the crux of evaluation.

- هذا هو جو هر التقييم.
- The actual organisation of the writeup of such an evaluation can be done several ways.
  - المنظمة الفعليه للكتابة حتى من هذا التقييم يمكن أن يتم بعدة طرق
- The most popular and sensible probably is to describe (b) fully in advance, and the relevant research/theory background to (c).
  - الأكثر شعبية ومعقول ربما هو وصف (b) بالكامل مقدما، وبحوث / نظرية المعلومات الأساسية ذات الصلة إلى (c).

- Then go through a systematic set of (a) points different aspects of the materials giving a clear description of each aspect and the actual evaluation (c) of each in relation to (b).
  - ثم انتقل من خلال مجموعة منتظمة من نقاط (a) الجوانب المختلفة للمواد إعطاء وصف واضح لكل جانب والتقييم الفعلي (c) من كل في ما يتعلق (b).
- Some people use the overt structure of the specific materials themselves as the (a) basis for proceeding. E.g. instead of having a prior idea of what categories to look at (e.g. from a published checklist), and using headings such as 'language content', 'balance of focus on the four skills' etc., they proceed through a list like 'reading passage', 'cloze exercises' (i.e. things the programmers present as separate parts of the materials).
- يستخدم بعض الناس هيكل علني للمواد المحددة نفسهم على أنهم (a) أساس للمضي قدما. مثلا بدلا من وجود فكرة مسبقة من فئات ما للنظر في (مثلا من نشر قائمة مرجعية)، واستخدام عناوين مثل "محتوى اللغة ' ميزان التركيز على المهارات الأربع 'وما إلى ذلك، أنها تنطلق من خلال قائمة مثل' ممر للقراءة ' ممر للتمارين '(أي تقديم الأشياء للمبرمجين كأجزاء منفصلة من المواد).
- That is in some ways 'easier' but of course instead of the evaluator imposing a relevant set of categories of things to look at it puts the materials in the driving seat and may mean that relevant things do not get looked at.
- وهذا هو لبعض الطرق 'الأسهل' ولكن بالطبع بدلا من مقيم فرض مجموعة ذات صلة من فئات الأشياء للنظر في أنه يضع المواد في مقعد القيادة، وربما تعني أن الأمور ذات صلة لا تحصل على النظر فيها.
- Compare what happens when you visit TESCO without a shopping list of one's own made in advance, and just uses the shelves of the store as a prompt for what to buy as one goes round!
  - قارن بين ما يحدث عند زيارة TESCO دون لائحة تسوق من المرء مقدما، ويستخدم فقط رفوف المخزن كما موجه لشراء واحدة وغني عن الجولة!
- Methods of evaluation (B): Empirical evaluation
  - طرق تقييم (B): التقييم التجريبي

- Other methods of evaluation generally require much more work, and for the materials to have been used for some time by learners/in actual classes (compare situation 3), so they are often firmly fixed in a specific teaching/learning situation (b).
- عموما غيرها من أساليب التقييم تتطلب الكثير من العمل، وبالنسبة للمواد أن تكون قد استخدمت لبعض الوقت من قبل المتعلمين / في الطبقات الفعلية (قارن الوضع 3)، لذلك غالبا ما تكون ثابتة بحزم في التدريس المحدد / وضع التعلم (b).
- However, they do move away from the purely introspective approach.
  - إلا أنها لا تبتعد عن نهج الاستقراء البحت.
- These are the ones that incorporate activities that are just like those we
  would otherwise regard as typical of regular empirical 'research' measurement, surveys etc. I.e. they may entail using questionnaires and
  interviews, systematically observing, eliciting 'think-aloud' data from
  software users, or testing users.
  - هذه هي تلك التي تتضمن الأنشطة التي هي تماما مثل تلك التي ستعتبر باسم نموذجية من العادية التجريبية 'البحث' القياس، الخ. اي أستطلاع أنها قد تنطوي على استخدام الاستبيانات والمقابلات، ومراقبة منتظمة، لأنتزاع البيانات "التفكير بصوت عال" من مستخدمي البرمجيات، أو مستخدمي الاختبار.
- They may mean doing 'studies' (experimental or not) comparing the success of one material against another and so forth, or indeed doing 'action research' with CALL. (See Chapelle, Jamieson and Park 1996 in ed. Pennington The Power of CALL for an overview of types of empirical research done on CALL classified by the kinds of methods used; and Chapelle 2001 pp66-94 for a more detailed coverage, in relation to CALL tasks of the more communicative type, and classic SLA research issues looked at in CALL)
  - ويمكن أن يعني به 'دراسات' (تجريبية أو لا) مقارنة لنجاح أحد المواد ضد آخر، وهكذا أو القيام بالفعل "البحث الإجرائي" مع CALL. إد قوة بنينجتون CALL لمحة عامة عن أنواع البحوث التجريبية القيام به على CALL المصنفة حسب أنواع الأساليب المستخدمة؛ وشابيل للحصول على تغطية أكثر تفصيلا، فيما يتعلق CALL مهام لنوع أكثر تواصلية، وبدا القضايا البحثية في CALL

- In themselves these 'research' type activities are non-evaluative, in the sense considered here (except action research).
  - في أنفسهم هذه الأنشطة لنوع 'البحث' غير قابلة للتقييم، بمعنى النظر هنا (باستثناء بحوث العمل).
- They are best seen as scientific means of gathering facts and testing hypotheses which can then either remain as cold statements of fact about what the effectiveness of the materials is or what people's opinions about them are, or be exploited for practical ends as part of an evaluation exercise - i.e. to make decisions like those described at the start.
- وتعتبر أنها أفضل وسيلة علمية لجمع الحقائق واختبار الفرضيات التي يمكن بعد ذلك تظل إما تصريحات اردة من حقيقة حول ما مدى فعالية المواد هو أو آراء الناس حول منهم، أو أن تستغل لتحقيق غايات عملية كجزء من ممارسة التقييم أي إلى اتخاذ قرارات مثل تلك التي وصفها في البداية.
- Examples are:

- الأمثلة هي:
- Doing a survey of teachers and/or learners who have used the material and finding out how they use it, their difficulties, attitudes to the interest and usefulness of the content, tasks etc.
  - بإجراء دراسة استقصائية للمعلمين و / أو المتعلمين الذين استخدموا المواد ومعرفة كيفية استخدامه، والصعوبات التي يواجهونها، والمواقف لمصلحة وفائدة المحتوى والمهام الخ
- Checklists can come in here again. E.g. one can base a
  questionnaire to users around the same set of (a) and (b) points
  that might otherwise be the points one asks oneself about in A
  above.
  - يمكن أن تأتي القوائم المرجعية هنا مرة أخرى. مثلا يمكن للمرء أن يبني استبيان للمستخدمين حول نفس المجموعة (a) و (b) النقاط التي قد يكون خلاف ذلك يسأل المرء نفسه في أعلاه.

- Observing a class using the program, taping and making systematic notes on their difficulties, actions, strategies, what they say, the teacher's involvement etc. Or one can ask learners to keep a diary of their reactions.
  - مراقبة الفئة باستخدام البرنامج والتسجيل وتدوين الملاحظات المنهجية على الصعوبات التي يواجهونها، وما يقولون إجراءات وأستراتيجيات ، تورط المعلم وما إلى ذلك أو يمكن للمرء أن يسأل المتعلمين للحفاظ على مذكرات لردود أفعالهم.
- Getting the computer to store records of actions performed by learners using a program and analysing them to infer learner strategies and processes. (E.g. revisions when wordprocessing, accesses made to an online glossary when reading).
  - الحصول على جهاز الكمبيوتر لتخزين سجلات الإجراءات التي يؤديها المتعلمين باستخدام برنامج وتحليله لاستنتاج الاستراتيجيات وعمليات المتعلم. (التنقيحات على سبيل المثال عند معالجة الكلمات، المنافذ التي قدمت إلى المعجم عبر الإنترنت عند القراءة).
- Example in T. Johns 1997 'Contexts' in ed Wichmann et al Teaching and Language Corpora (Longman).
  - مثلاً في T. جونز 1997 اسياقات في إد يشمان وآخرون لتدريس اللغة والمجاميع (لونجمان).
- The classic research comparison of those using one program
  with those using another differing in a small or large way (or no
  program... just doing non-computer equivalent tasks) over a
  period, with before and after tests to check on how much has
  been learnt.
- مقارنة البحوث الكلاسيكية من تلك التي تستخدم برنامج أحد أولئك الذين يستخدمون اختلاف آخر في وسيلة صغيرة أو كبيرة (أو أي برنامج ... مجرد القيام بمهام مماثلة غير الكمبيوتر) على مدى الفترة، مع قبل وبعد الاختبارات للاطمئنان على مدى تم تعلمه .

- If A type and B type evaluation are both done, the connection between the two needs to be spelt out.
  - إذا كان نوع وتقييم نوع B على حد سواء القيام به، والعلاقة بين احتياجات اثنين إلى تهجئتها
- If the A evaluation resulted in adoption of the software, did the B evaluation show that was a good decision?
  - إذا نتج عن تقييم A في اعتماد البرمجيات، لم تظهر أن تقييم B كان قرارا جيدا؟