MEANDRIX

A MEANDRIX PRIMER

# Sequenced Dialogic Assessment

*Using personalised evidence to assess insight, growth and application.*

AUTHOR

Dr. Gary Crowfoot Ph.D. (Nursing) SFHEA

Founder, Meandrix Education Systems

SECTION 01

# Why assessment must do more than check answers

Assessment in higher education has often rested on the artefact generated by a student to demonstrate their knowledge. Some of the classics include the essay, the report, the exam paper. In my field of nursing, we also have the classic written case study. The idea is the generation of the artefact itself tells us the learning journey the student has been on to arrive at the submitted product. Yet, we must now consider if we have reached the limit of the standalone artefact in the modern age of education.

Please do not mistake my intention. Artefacts remain useful. They represent what a student produced. They can be moderated, marked, archived, and compared. They can be delivered at scale and assessed at the marker's convenience without requiring the presence of the student. However, artefacts have always had limits. Now, those limits have been tested, stressed, and are impossible to ignore. A submitted product does not always show how the student arrived at it. It does not reveal which decisions were considered and rejected, which were taken without thought, what feedback was acted on, or how the student would handle the same situation in different circumstances. The artefact captures the end result. It does not always capture reasoning, growth, or knowledge application the student went through to achieve it. Especially now that Generative AI has shortened the distance between idea and the polished product.

Universities and higher education have responded to this challenge by increasing activities that many educators now recognise, or experience, as academic policing. From plagiarism checkers that scan for AI use to a return to the formal invigilated exam, there is movement in the sector to secure assessment. However, there is more than one way to secure an assessment. Perhaps, we need to think about securing assessment through the way we design it. We might even just be able to provide richer evidence of learning along the way.

SECTION 02

# Where this primer comes from

I started thinking about this problem in parallel with starting Meandrix. Over my time from a Registered Nurse on the hospital floor to convening a large Bachelor of Nursing program, I have routinely encountered the student who can write the perfect essay but struggles to deliver care. I have also witnessed the inverse where the knowledge could be demonstrated, but could not be translated and transcribed to the written task. In each case, the artefact often missed the most important thing: the reasoning behind the actions taken.

Recently I was reading *The Mom Test*. It is a short book by Rob Fitzpatrick designed to help would-be entrepreneurs engage in customer discovery conversations. Its central argument is simple. If you want to learn whether a problem is real, do not ask people whether they would buy your product or whether your idea is good. Ask them what they have already done, in real life, about that problem.

The principle stayed with me. Observed action is better evidence than hypothetical intention. I believe that the same idea can apply in our approach to assessment. If you want to understand a student's clinical reasoning, do not start by asking how they would respond to a hypothetical scenario. Take it back one step. Start with them completing a task or action in a controlled way. Then, ask the hard question. Ask why. Why is a question that drives them to delve beneath their decisions, explain their reasoning, explore what they learned, and evaluate the outcome. It takes them beyond "because that is what I should do" to "this is the logic behind my actions."

Two ideas from very different fields. The same underlying logic. This primer is an attempt to put that logic into an assessment design that holds up academically.

SECTION 03

# Sequenced dialogic assessment: a working definition

Dialogic assessment has a long history in higher education. The viva voce, the thesis defence, the presentation, the interview — these are familiar forms. At their best, they are conversations informed by a body of evidence the student has already produced. The thesis defence is the apex example: years of research and writing made auditable through a verbal interrogation.

Generative AI has now made it possible to produce a presentation in seconds, delivered by an AI avatar that looks and sounds like the learner. The form looks dialogic. The evidence sequence behind it is absent. Both the thesis defence and the AI-generated presentation are technically dialogic. Only one rests on a real learning journey.

Sequenced dialogic assessment is the pattern that distinguishes these two. It requires a real sequence of evidence before the conversation begins. The thesis defence is the apex of this pattern at program scale. What this primer describes is the same pattern at unit scale — designed so that an individual educator can build it inside their own subject.

More formally: sequenced dialogic assessment is a unit-level assessment design in which students complete connected assessment tasks that progressively build evidence of learning, followed by an interactive assessment conversation that uses that evidence to explore the student's reasoning, reflection, growth and application. It is not a single instrument. It is a design pattern.

The label is not new in spirit. Connected assessment has been conducted for many years. This is simply the term I am using to describe the design pattern in this primer. The design borrows from three strands of existing assessment scholarship.

First, **programmatic assessment**, developed by van der Vleuten et al. (2012), argues that high-stakes judgements should be based on patterns of evidence collected over time rather than single performances — repeated assessment of performance rather than point-in-time observation. The point is to allow safe formative learning while enabling appropriate high-stakes summative assessments. Sequenced dialogic assessment is an attempt to apply this principle at unit level rather than program level.

Second, **dialogic feedback**, associated with the work of Carless and Boud (2018) and Ajjawi and Boud (2017), treats feedback as a conversation that develops feedback literacy and supports learning. This interrogation of understanding allows surfacing of the underlying reasoning the learner presents, rather than producing a one-way report on a finished product. I like to think of this as the

photograph and the movie. The photograph represents traditional feedback: expert driven, captures a point in time, and can only use the evidence presented. For some, the image is clear. For others, it is a blurry photograph taken by a toddler. The dialogic feedback represents a movie. You can interrogate the plot, surface themes and understanding, examine reasons and actions, and allow a more nuanced evaluation. Sequenced dialogic assessment attempts to connect assessment and feedback in a way that gains greater insight into the capabilities of the learner.

Third, **interactive oral assessment**, explored by Joughin (1998), has seen a widespread revival in response to the surge in accessibility and use of generative AI. This approach asks students to speak to their work rather than only submit it. For its application to sequenced dialogic assessment, the most important concept is informing the conversation through prior evidence rather than treating it as a free-standing oral exam.

To be clear, this is not full programmatic assessment. Programmatic assessment is conceptualised at the program level. What this primer describes is a design that uses selected programmatic principles within a single unit. The distinction matters. Programmatic assessment, properly implemented, requires institutional commitment that most unit conveners cannot command alone. Sequenced dialogic assessment is what an individual educator can build inside their own subject.

SECTION 04

# How this differs from presentation-plus-questions

One common response to AI in assessment has been to ask students to present their work and answer questions. This is a reasonable step. It requires the student to speak to their own submission and gives the assessor some assurance about authorship. Yet, it has two quiet weaknesses.

The first weakness is exposed when the follow-up questions are generic or randomised. Suddenly the conversation is depersonalised, static, and returns to the recall or rote learning approach. The student may answer correctly, fluently, and confidently without demonstrating any insight into their own reasoning. A polished speaker can survive a presentation with surface-level engagement. A nervous expert can be derailed by an unlucky question. Randomised questions can make assessment procedurally fairer. They cannot necessarily make it valuable for both learner and educator.

The fairness of the approach does not depend on every student receiving identical questions. It depends on every student being assessed through a common questioning framework, with probes anchored in evidence from their own demonstrated decisions.

The other weakness is scale. If you refuse to make the above compromise for the purpose of scale, then you challenge markers with actively recalling what a student reported, identifying the points of exploration, and formulating them into questions in the moment. This can be strong for exposing decision and reasoning. Yet, the cost in workload, time, and capability makes scaling this approach close to impossible.

Sequenced dialogic assessment proposes a different starting point. The personalisation does not depend on the assessor's recall in the moment. It is built into the design of the assessment itself. The conversation begins with evidence of what the student has actually done, captured during prior assessment tasks. The questions are not drawn from a pool. They are anchored in the student's own pathway of decisions and actions. This narrows the conversation and deepens it.

SECTION 05

# A questioning logic anchored in action

The questioning logic that runs through this design starts in the same place as the Mom Test. Observed behaviour first. Reasoning second. Learning third. Transfer last.

In practical terms, the assessor moves through five layers.

1. What did you do?
2. Why did you do it?
3. What did you learn?
4. What would you do next?
5. Why?

The most important question is *why*. Insight is usually revealed in the reasoning behind the action, not in the action itself.

The order also matters. The move to "what would you do next" should only happen once past behaviour has been thoroughly established. Asking about future action before the actual decisions have been explored invites the rehearsed, hypothetical answer the principle was meant to avoid.

This logic does not require theatrical questioning. It requires evidence. The assessor needs to know what the student did. Without that, every question becomes hypothetical, and hypothetical questions invite rehearsed answers.

SECTION 06

# Three honest challenges

Sequenced dialogic assessment is not free of difficulty. There are three challenges that any honest account has to name.

**Personalisation.** Dialogic assessment is strongest when it is anchored in the student's own work. That requires access to individual evidence. Without it, questions collapse into generic prompts that test fluency rather than insight. This is a real constraint, not a theoretical one. Most assessors do not have time to reconstruct each student's decision pathway by hand.

**Scalability.** Dialogic assessment takes time. As cohort size increases, it becomes harder to maintain the depth, consistency and responsiveness that make the conversation valuable. The earlier argument applies here too. Personalisation that depends on the marker's memory does not scale. What changes in sequenced dialogic assessment is that the personalisation is embedded in the design. The evidence is produced before the conversation begins, so the marker is not building it from scratch. This makes the approach more scalable than what it replaces, but not infinitely so. The conversation itself still takes time. Squeezed too hard, it either becomes unsustainably labour-intensive or it is simplified into a shallow oral check that does not deserve the name.

**Integrity.** Dialogic assessment can strengthen integrity, but integrity is not automatic. A conversation only strengthens integrity if it is anchored in the student's own work. Asking students to speak in general terms about a topic is a much weaker signal than asking them to explain specific decisions they have already made. The integrity value lies in the evidence, not in the act of speaking.

**Defensibility.** For the approach to be defensible, the conversation must still be mapped to stated learning outcomes, supported by transparent criteria, and moderated through consistent assessor judgement. Personalisation should guide the questioning but cannot replace fairness, calibration or criterion-based marking.

Naming these challenges matters. A primer that pretends they are solved is not a primer worth writing.

SECTION 07

# Where Meandrix fits

Meandrix was not designed from the outset as a dialogic assessment platform. It was designed for applied scenario-based assessment, where students make decisions inside simulated situations and the platform captures what they choose, what consequences follow, and how they respond to changing information.

In doing this, Meandrix produces exactly the kind of evidence that a sequenced dialogic assessment design depends on. Each student's pathway is captured. Decisions, alternatives, and consequences are recorded. From that record, a personalised report can be produced.

It is worth saying clearly that Meandrix is one implementation of this pattern, not the only path. A paper-based portfolio, a multi-stage assignment, or a structured reflection log could carry the evidence sequence in principle. What Meandrix does is automate the capture and produce the personalised report at the scale a unit usually requires.

That report is not the evaluation. It is a conversation guide. It supports, but does not replace, professional judgement. The report tells the assessor where to probe, what to ask the student to explain, and where reasoning developed or deteriorated. The educator remains central. The marking is theirs. The evaluation is theirs.

It is worth being honest about how Meandrix addresses the three challenges, because it does not address them evenly.

It directly addresses personalisation. The personalised report is the point.

It contributes to integrity, because the conversation can be anchored in decisions the student has demonstrably made, not in general claims about a topic. Put another way: this is integrity through assessment design rather than surveillance. A task that requires real reasoning, anchored in evidence the student has visibly produced, is harder to substitute than it is to police.

It partially improves scalability. The conversation itself still takes time, but the preparation is faster and the evidence is more reliable. In a large cohort, where multiple assessors share the marking load, a personalised report grounded in the student's own decisions is a better starting point than another marker's notes or a partial impression of the work. The conversation does not get shorter, but it gets better at scale.

That is the most honest framing I can offer. Meandrix does not replace the educator's work. It changes the quality of the evidence the educator has when they do that work.

SECTION 08

# From activity to evidence to conversation to evaluation

The design pattern can be summarised in four stages.

| STAGE | WHAT HAPPENS | THE ROLE OF EVIDENCE |
|---|---|---|
| **Activity** | The student completes an applied decision-making task | Decisions, pathways and consequences are captured |
| **Evidence** | The activity is summarised into a personalised record | A report of individual effort is generated |
| **Conversation** | The educator conducts a dialogic assessment | The report guides targeted questioning |
| **Evaluation** | The educator assesses insight, growth and application | The evaluation is informed, defensible and traceable |

From activity to evidence. From evidence to conversation. From conversation to evaluation.

SECTION 09

# A practical questioning framework

| LAYER | PURPOSE | EXAMPLE PROMPTS |
|---|---|---|
| **Describe** | Establish what the student did | "Talk me through what happened at this point." |
| **Explain** | Explore reasoning | "Why did you choose that action?" |
| **Analyse** | Examine feedback, evidence and alternatives | "What else could you have considered?" |
| **Evaluate** | Judge growth and learning | "Looking back, what does this show about your development?" |
| **Apply** | Transfer learning to future practice | "What would you do next time, and why?" |

The framework is sequential but not rigid. An assessor may move back and forth between layers, but the Apply layer should not be reached until the earlier layers have done their work. Asking what the student would do next before establishing what they actually did invites speculation rather than evidence.

SECTION 10 · RUBRIC

# A model rubric in development

The rubric below is designed for the dialogic component of a sequenced assessment. It should be used alongside the evidence generated through the preceding task, rather than as a replacement for judging the quality of the original performance. It is offered as a starting point rather than a finalised instrument. It will need further development, local context, discipline nomenclature and modification to meet institutional requirements. Students should not be assessed on verbal confidence, fluency or presentation style. The focus is on the quality of reasoning, the accuracy of action reporting, the explanation of initial decisions, the analysis and evaluation of learning, the proposed future action, and the rationale that supports it.

| CRITERION | HIGH DISTINCTION (8.5–10) | DISTINCTION (7.5–8.4) | CREDIT (6.5–7.4) | PASS (5.0–6.4) | FAIL (0–4.9) |
|---|---|---|---|---|---|
| **1. Report actions taken** *(What did you do?)* | Comprehensive, accurate and precise account of all relevant actions and decisions, with an appropriate level of detail. Clearly identifies and organises the sequence of events using the available evidence. | Clear, accurate account of most relevant actions and decisions, with only minor omissions or variation in detail. | Sound account of key actions, though some elements may lack precision, detail, or completeness. | Basic account with notable omissions, vague descriptions, or some inaccuracy. | Fragmentary, inaccurate, or substantially incomplete account of what was done. |
| **2. Explain initial reasoning** *(Why did you do it?)* | Sophisticated explanation of initial reasoning that integrates context, evidence, and disciplinary expectations. Shows clear self-awareness of the logic that drove the original decisions. | Clear explanation with sound reasoning and confident engagement with disciplinary context. | Adequate explanation showing basic reasoning, with some links to disciplinary context. | Basic explanation, largely descriptive, with limited articulation of reasoning or weak disciplinary grounding. | Lacks coherent explanation of initial reasoning, or explanation is inaccurate or absent. |
| **3. Analyse and evaluate learning** *(What did you learn?)* | Critically analyses what was learned with strong examination of feedback, alternatives, and consequences. Demonstrates well-supported evaluative judgement of growth and meta-cognitive awareness. | Effective analysis of learning with sound evaluation of growth. Engages substantively with feedback received. | Sound analysis of learning with adequate evaluation of growth, though depth may vary across elements. | Basic identification of learning, more descriptive than analytical or evaluative. Limited engagement with feedback. | Little or no meaningful analysis or evaluation of learning. |
| **4. Propose future action** *(What would you do next?)* | Comprehensive, specific, and contextually appropriate revised approach. Demonstrates clear forward thinking and the ability to translate learning into actionable practice. Identifies alternatives where relevant. | Clear, specific revised approach grounded in the learning sequence. Demonstrates sound forward thinking. | Appropriate revised action proposed, though specificity or contextualisation may be uneven. | Basic proposed action; generic, underdeveloped, or only loosely tied to the learning sequence. | Little or no meaningful proposal for future action. |
| **5. Justify revised reasoning** *(Why?)* | Sophisticated justification of the revised approach that integrates context, evidence, professional judgement, and capacity to transfer learning to novel contexts. Demonstrates strong second-order reasoning. | Clear justification with sound reasoning and confident engagement with disciplinary expectations. | Adequate justification showing basic reasoning behind the revised approach. | Basic justification; reasoning is generic or weakly developed. | Little or no coherent justification of the revised approach. |

SECTION 11

## A note of caution

Technology does not make dialogic assessment inherently valid, scalable or fair. Its value depends on how well it supports purposeful assessment design, consistent assessor judgement and meaningful student engagement.

Meandrix can make a dialogic conversation better informed. It cannot make it more important than it already is.

SECTION 12

## Closing position

Sequenced dialogic assessment offers a way to make learning more visible. It shifts assessment from a static evaluation of a final artefact toward an evidence-informed conversation about action, reasoning, growth and application.

The design is not new. It draws on programmatic assessment, dialogic feedback, and interactive oral assessment scholarship. What may be new is the willingness to put these strands together in a single unit-level design that an individual academic can use inside their own subject.

Generative AI is the catalyst, not the cause. The deeper reason to consider this design is that recall has never been the same thing as understanding, and the artefact has never been the same thing as the reasoning behind it.

Meandrix exists to help educators see that reasoning.

*Gary Crowfoot*

Founder, Meandrix Education Systems

# References


Ajjawi, R., & Boud, D. (2017). Researching feedback dialogue: An interactional analysis approach. *Assessment & Evaluation in Higher Education, 42*(2), 252–265. https://doi.org/10.1080/02602938.2015.1102863

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Longman.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Fitzpatrick, R. (2013). *The Mom Test: How to talk to customers and learn if your business is a good idea when everyone is lying to you.* Founder Centric.

Joughin, G. (1998). Dimensions of oral assessment. *Assessment & Evaluation in Higher Education, 23*(4), 367–378. https://doi.org/10.1080/0260293980230404

van der Vleuten, C. P. M., Schuwirth, L. W. T., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K. J., & van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher, 34*(3), 205–214. https://doi.org/10.3109/0142159X.2012.652239