Calibrating Assessment Workload and Task Design – Quick Reference Guide

Key Principle:
Assessment workload should reflect the effort required for students to think, decide, and justify—not the length, format, or duration of the task.
It is determined by cognitive complexity and decision-making demand, shaped by task constraints set by the assessor.
This guide provides a consistent approach to designing and calibrating assessment workload using these factors as the primary determinants of effort.

The following definitions describe the core elements used to design and calibrate assessment workload:

Key Definitions

Cognitive effort	The level of thinking required expressed through observable and measurable actions (e.g., identify, apply, analyse, evaluate, justify, design, create).
Decision-making demand	The extent to which students determine, structure, and justify their approach. • **Directed** – Decisions are predetermined; students follow instructions or select from fixed options • **Guided** – Some choice within a defined structure; limited independent judgement required • **Self-directed** – Students determine approach, structure, and focus • **Evaluative / Justified** – Students must justify decisions, evaluate alternatives, and defend their reasoning
Constraint profile	Indicates the **conditions set by the assessor that shape how thinking and decisions must be performed**. Constraints may increase or reduce workload depending on how they affect precision, scope, and justification. • **Directive** – Highly specified conditions that reduce scope and limit decision-making • **Bounded** – Defined limits (e.g., format, length) with some scope for interpretation • **Open** – Minimal constraints; students define scope and approach • **High-compression** – Tight limits requiring prioritisation, precision, and selective decision-making • **Adaptive / Performance** – Real-time conditions requiring responsive thinking and justification
Artefact	Any output or evidence produced by a student that demonstrates learning, including both final products and relevant evidence of the process used to create them. Where appropriate, targeted artefacts may be used to make student thinking, decision-making, and iteration visible (e.g., plans, drafts, rationales, decision-making strategy). Such artefacts should be purposeful and proportionate, strengthening evidence of learning without introducing unnecessary or duplicative workload.

Reframing Common Assumptions About Workload

The following table highlights common assumptions that can lead to misalignment between assessment design and actual student workload. These are reframed to align with the principles outlined above.

Common Assumption	Impact on Practice	Reframed Position
Word count reflects workload	May overestimate effort for descriptive tasks and underestimate effort for analytical or evaluative tasks	Workload should be calibrated based on cognitive complexity and decision-making demand, not output length
Longer presentations require more effort	May undervalue short, high-compression tasks that require prioritisation and precision	Shorter tasks may require greater effort due to selectivity, clarity, and audience adaptation
Output volume indicates student effort	AI tools can generate large volumes of content quickly, obscuring actual student contribution	Effort should be evidenced through judgement, selection, and justification, not volume of output
Tasks of similar length have similar workload	Masks variation in cognitive complexity and decision-making demand	Tasks of equal length may differ significantly in effort depending on the level of thinking and judgement required
Constraints reduce workload by limiting output	May overlook increased demands for precision, prioritisation, and justification	Constraints shape workload and may increase or reduce effort depending on how they influence task requirements

Using the Tiered Tables to Calibrate Assessment Workload

The tiered tables support staff to **scope and calibrate assessment tasks** based on the level of thinking required, the extent of student decision-making, and the conditions under which the task is performed.

Assessment workload (i.e., tier level) can be adjusted without changing the task format by modifying:

- **Cognitive complexity** (e.g., understanding → analysing → evaluating → creating)
- **Decision-making demand** (e.g., directed → guided → self-directed → justified)
- **Task constraints** (such as word count, time, format, or performance requirements) are conditions that define how a task is completed. They guide precision, prioritisation, and justification but do not directly measure workload.

Workload is determined by how much students must think, decide, and justify—not by the length, format, or duration of the task.

Applying the Tables

Use the tables to scope and calibrate assessment tasks as follows:

1. **Identify the closest assessment type** *(examples are illustrative, not exhaustive)*
2. **Select the appropriate tier** based on cognitive complexity and decision-making required
3. **Use the indicative weighting and hours as a guide**
   Indicative weightings and hours are ranges and should be interpreted and adjusted based on cognitive complexity, decision-making demand, constraint profile, and artefact requirements.
4. **Adjust the task as needed** by modifying:
   - cognitive complexity
   - decision-making demand (level of independence and justification)
   - task constraints, such as scope, time, format, and performance conditions, shape how students carry out a task. When constraints involve word limits or duration limits, these define how the task should be performed rather than the amount of effort required.
   - number and type of artefacts

Where tasks include multiple artefacts, **distribute workload across artefacts based on the cognitive effort required for each**, rather than the size or prominence of the final output.

Avoid combining multiple high-effort (Tier 1) tasks within a single unit unless justified by credit point value and overall workload allocation. This will prevent over-assessment and strengthen quality assurance and defensibility. Apply the same principles to other or discipline-specific assessment types. Tasks should be

calibrated so that workload reflects the effort required to think, decide, and justify, rather than the volume of output produced.

Over-reliance on Tier 3 tasks for summative assessment may limit opportunities to evidence higher-order thinking. These tasks are most effective for formative, early-stage, or low-stakes use, and should be balanced with Tier 1 and Tier 2 tasks that better assess judgement and application.

Assessment tasks across all tiers can be designed within either of the dual lanes (independent capability or AI-integrated) of the FedUni ASSURE framework, with expectations for human judgement, transparency, and appropriate use of tools made explicit through task design and artefact requirements.

Alignment Requirements

Ensure all assessment tasks:

- align with the **AQF level** of the unit
- fit within the **total unit workload**:
    - **15 credit points ≈ 150 hours**
    - **30 credit points ≈ 300 hours**
- contribute to a balanced assessment profile across the unit with total assessment workload proportionate to overall study expectations

---

The tiered tables provide indicative examples of how assessment tasks can be classified and calibrated based on cognitive effort, decision-making, and constraints.

Tier 1 – High Cognitive Effort (Approx. 30–50 Hours per Task)

High independence, high judgement, and iterative or integrative work requiring justification of decisions. Produces strong evidence of student thinking, reasoning, and process.

Assessment Type	Typical Output	Cognitive Effort	Decision-Making Demand	Constraint Profile	Indicative Weighting	Indicative Hours	Design Recommendations	Artefacts
Project / Design Task	Artefact/s + docs	Design, develop, refine, justify	Self-directed → Evaluative	Open (iterative; staged deliverables)	40–60%	35–50 hrs	Assess process & iteration	Logs, version history
Portfolio	Mixed	Curate, synthesise, evaluate	Self-directed → Evaluative	Composite (multi-component artefacts)	40–60%	30–45 hrs	Align artefacts	Justified artefact selection
AI Task	Prompt + critique	Evaluate, refine, justify	Evaluative	Tool-mediated (bounded; iterative prompting cycles)	25–45%	20–35 hrs	Assess judgement	Prompt logs, critique
Live Oral / Viva	Live Q&A	Explain, defend, evaluate	Evaluative / Justified	Adaptive (performance; live questioning)	30–50%	20–35 hrs	Assess reasoning	Performance
Simulation / Scenario Task	Decision + rationale	Analyse, decide, justify, adapt	Self-directed → Evaluative	Adaptive (time-bound; scenario constraints)	30–50%	20–35 hrs	Use authentic scenarios	Decision log, rationale
Capstone / Integrated Task	Multi-component	Integrate, evaluate, justify, create	Self-directed → Evaluative	Open (composite; multi-artefact integration)	40–60%	40–50+ hrs	Assess integration	Synthesis commentary
Research / Inquiry Task	Proposal / report	Investigate, evaluate, synthesise	Self-directed	Open (independent inquiry scope)	30–50%	25–40 hrs	Emphasise question design	Research plan, source rationale
Pre-recorded Video / Digital Presentation	Short-form video (e.g. 1–3 min) + supporting artefacts	Integrate, evaluate, justify, translate, refine	Self-directed → Evaluative / Justified	High-compression (multimodal; e.g., 1–3 min video as a performance constraint requiring prioritisation and justification)	30–45%	25–40 hrs	Assess integration of knowledge, strategic communication choices, and justification—not production polish	Storyboard, draft iteration, decision rationale, audience justification, (optional) AI-use critique

Tier 2 – Moderate to High Effort (Approx. 15–30 Hours per Task)

Moderate independence and judgement, with structured guidance. Emphasises analysis, application, and some justification, with partial visibility of process.

Assessment Type	Typical Output	Cognitive Effort (Observable Verbs)	Decision-Making Demand	Constraint Profile	Indicative Weighting	Indicative Hours	Design Recommendations	Artefacts
Critical Analysis	Written	Analyse, evaluate	Self-directed	Open (evaluative; extended written response)	30–45%	20–30 hrs	Focus on judgement	Annotated outline
Report	Structured	Analyse, synthesise	Guided → Self-directed	Structured (bounded; e.g., word limit or template)	30–50%	20–30 hrs	Assess synthesis	Draft rationale
Presentation	Slides	Select, explain	Self-directed	High-compression (e.g., 3–5 slides or 3–5 minutes as a constraint requiring selective communication and prioritisation)	25–40%	15–25 hrs	Assess clarity	Slide annotations
Case-Based Decision Task	Decision + brief	Apply, analyse, justify	Guided → Self-directed	Bounded (applied; case parameters constrain scope)	25–40%	15–25 hrs	Focus on application	Decision rationale
Problem Set (Open-ended)	Worked solutions	Apply, analyse, justify	Guided → Self-directed	Structured (interpretive; non-routine problems within defined parameters)	20–35%	12–24 hrs	Include non-routine problems	Working steps
Peer Review / Evaluation	Feedback	Evaluate, critique, justify	Evaluative	Bounded (criteria-based judgement)	10–25%	8–15 hrs	Assess quality of judgement	Annotated feedback
Lab / Practical	Execution + report	Apply, analyse, interpret	Guided → Evaluative	Procedural (interpretive; controlled lab conditions)	25–40%	15–25 hrs	Assess interpretation	Lab notes, logs
Pre-recorded Video (Short-form)	Video (e.g., 3–5 min)	Analyse, synthesise, translate, explain	Self-directed → Evaluative	High-compression (multimodal; e.g., short video requiring prioritisation)	25–40%	15–30 hrs	Assess clarity, prioritisation, audience adaptation	Script/storyboard, rationale

Tier 3 – Foundational Effort (Approx. 6–15 Hours per Task)

Guided and structured tasks with lower decision-making demand. Focuses on understanding, interpretation, and initial application.

These tasks are most effective for formative, early-stage, or low-stakes use, and should be balanced with Tier 1 and Tier 2 tasks that better assess judgement and application

Assessment Type	Typical Output	Cognitive Effort (Observable Verbs)	Decision-Making Demand	Constraint Profile	Indicative Weighting	Indicative Hours	Design Recommendations	Artefacts
Short Response	Written	Identify, explain	Guided	Bounded (e.g., 150–300 words as a scope constraint requiring selectivity and precision)	15–25%	8–15 hrs	Require interpretation	Brief reasoning
Quiz	Timed	Recall, identify	Directed	Directive (timed; fixed responses)	10–20%	6–10 hrs	Use for checks	None
Annotated Bibliography	Source summaries	Identify, evaluate, summarise	Guided	Bounded (e.g., defined number of sources)	15–25%	10–18 hrs	Require evaluation	Source rationale
Concept Explanation	Written/oral	Explain, translate, simplify	Guided	Bounded (short explanation for defined audience)	15–25%	8–15 hrs	Target non-expert audience	Explanation rationale
Structured Reflection	Written/audio	Describe, explain, reflect	Guided	Directive (prompted; structured reflection format)	10–20%	6–12 hrs	Use targeted prompts	Evidence-linked reflection

Statement of AI Use

This document was developed with the assistance of generative artificial intelligence to support drafting, synthesis, and language refinement. The authors have exercised full oversight, applying professional judgement to review, verify, and contextualise all content in accordance with Federation University's requirements and academic standards.