

Core Learning Model — Validation Plan

The Grow Space · a staged plan to establish the psychometric credibility of the tool

Purpose. This sets out how we move the Core Learning Model from a well-designed prototype to an **evidence-backed instrument**, beginning with a free school pilot. It names the studies, the analyses, the sample sizes and the outcome measures — the document we would take to a university partner / Human Research Ethics Committee (HREC). It is deliberately staged so each phase produces usable evidence on its own.

1 Foundations to put in place first

- **Research consent.** A separate, opt-in consent ('use my de-identified responses to research & improve the tool'), stored server-side with its own version tag. Validation may only use data collected under this consent. (*Being built now.*)
- **Ethics approval.** Ideally via a **university partnership** — gives HREC cover, academic co-investigators and independence. The single biggest credibility lever.
- **Data capture.** Persist item-level responses, reaction times and timestamps (not just computed scores) — required for reliability and factor analysis. A de-identified **cohort export** (service-role) feeds the analysis pipeline (Python/R).
- **Pre-registration.** Register the analysis plan (e.g. OSF) before data collection to signal rigour.

2 What we are validating

Each module is a measurement claim to be tested independently:

- **Cognitive screens** — working memory (forward/backward span), processing speed, reasoning, verbal fluency, visual/auditory channel.
- **Self-report scales** — Big Five (personality), RIASEC (interests), VIA-style character strengths, study habits / metacognition / self-regulation / environment.
- **Derived outputs** — the banding, the strengths narrative, the subject/careers suggestions, and the 3-session program's effect.

3 The validation framework

Property	Question it answers	Method	Min. n (guide)
Internal consistency	Do items on a scale hang together?	Cronbach's α / McDonald's ω ; item-total correlations	~100+
Test-retest	Is it stable over time?	Re-administer to a subset @ 2–4 wks; correlate	~40–60 subset
Factor structure	Does it measure the constructs it claims?	EFA then CFA; model fit indices	~200–300
Convergent / discriminant	Does it agree with gold-standard measures?	Run an established measure on a subsample; correlate	~60–100 subset
Norms	What is 'typical' for this age?	Percentile tables by year/age from a representative sample	500–1000+
Criterion / predictive	Does it predict real outcomes?	Link to grades / teacher ratings; follow-up	200+ longitudinal
Incremental	Does it add value beyond prior GPA?	Hierarchical regression over existing predictors	200+
Consequential / efficacy	Does acting on it improve learning?	Quasi-experiment / small RCT vs control	2 cohorts
Fairness	Is it equitable across groups?	Differential item functioning (gender, EAL, SES)	300+

4 The phased plan

Phase 1 — Single-school pilot (this term; n ≈ 100–300)

- **Delivers:** internal consistency for every scale; item analysis (revise/cut weak items); preliminary factor structure; completion & drop-off analytics; time-on-task; face-validity feedback from students & teachers.
- **Plus:** a test-retest sub-sample (40–60 students re-take @ 3 weeks).
- **Outcome:** a v2 instrument with the noisy items removed, and the first internal evidence pack.

Phase 2 — Multi-school (2–3 terms; n ≈ 500–1000+)

- **Delivers:** confirmatory factor analysis; convergent validity vs gold-standard measures on a subsample; the first **norm tables** (turning relative bands into true percentiles); fairness / DIF analysis.
- **Outcome:** a Technical Manual draft and a defensible, normed scoring model.

Phase 3 — Longitudinal & efficacy (6–18 months)

- **Delivers:** predictive & incremental validity (profiles vs later grades / engagement, beyond prior GPA); a quasi-experiment or small RCT on whether the 3-session program improves outcomes vs a comparison group.
- **Outcome:** a peer-reviewed publication and the strongest evidence claim in this market.

5 Outcome measures for predictive validity

- School-provided **achievement** (report-card grades, internal assessment marks) — with consent and a data-sharing agreement.
- **Teacher ratings** of engagement / study skills (brief standardised form).
- Student **self-report follow-up** (did the strategies help? confidence, wellbeing) at 6–12 weeks.
- Program **engagement** metrics already captured (sessions completed, activities done).

6 How we increase credibility (highest-leverage first)

- **University partnership + HREC approval** — ethics cover, rigour, independence, a co-authored study, and a student-researcher pipeline.
- **Technical Manual / validation white paper** — the accreditation-substitute a school's psychologist will ask for.
- **Pre-registration** of the analysis plan; an **academic advisory board**; and ultimately a **peer-reviewed publication**.

7 Risks & how we manage them

- **Consent scope** — validation may only use data gathered under explicit research consent (see the Consent & Privacy brief). Built first.
- **Representativeness** — one school's cohort is not a norm sample; norms wait for Phase 2's multi-school data.
- **Self-selection & missing data** — track drop-off; report completion transparently rather than silently dropping cases.
- **Over-claiming** — until norms exist, results stay framed as relative-strength screening, reviewed by a psychologist — never a standardised diagnosis.

Immediate next steps: (1) ship research-consent capture; (2) stand up the de-identified cohort export; (3) approach a university partner with this plan to secure HREC approval; (4) run the Phase 1 pilot. Contact: greg@thegrowspace.com.au.