DISCUSSION PAPER · NOVEMBER 2026 EDITION

The PrimarAI School

A manifesto and design proposal for an AI-native Australian primary school

We have national principles, international experiments and a working tool in our own classrooms, but no worked example of what a primary school designed for this moment would actually look like in NSW

Dane Smale

NESA-accredited primary teacher and founder, PrimarAI

Sydney, November 2026

primarai.com.au

Published openly under a Creative Commons Attribution 4.0 licence

Contents

Preamble: Can you be mean to a chatbot? 7
Why this paper exists 7
The argument 8
How the paper is organised 9

Part I: Facing the future 10
1 What has changed 10
2 What schooling has achieved, and what it assumes 15
3 Three responses, and what they leave open 16
4 Why integration is not enough 21

Part II: What is primary education for? 23
5 Education for now, not only for later 23
6 Foundations, old and new 24
7 What a primary school should cultivate 25
8 The value of friction 26
9 What must remain human 28

Part III: The principles of an AI-native school 30

Part IV: The PrimarAI day 35
10 The ARC Day 35
11 Augment 40
12 Reason 49
13 Digital literacy 51
14 Create and the five Cs 53
15 How the parts make one day 56
16 The school year 57
17 An example Wednesday at the school 57

Part V: The school in practice 60
18 Curriculum and the NSW syllabuses 60
19 Progression, grouping and home groups 65
20 Teachers 66
21 Assessment in six layers 68
22 Graduation commitments 70
23 Families, community and partnerships 72
24 Wellbeing, inclusion and safeguarding 72
25 Governance, privacy and data 73
26 Place, community and an open model 74

Part VI: Evidence, uncertainty and implementation 76
27 Five kinds of claim 76
28 What the evidence says 77
29 Objections and responses 79
30 Testing the model 81
31 What would count against the model 82
32 Implementation and the conditions for success 83

Epilogue: An invitation to build the next school 84

References 86

Appendix A: Two NSW units taught the ARC way 101
A.1 English, Stage 1: Unit 20, Colour Me 101
A.2 Mathematics, Stage 2: Unit 20 104
A.3 The pilot coaching course 107

Abstract

Artificial intelligence is now part of the conditions in which Australian students learn. In NSW public schools a purpose-built chatbot is available from Year 5, and outside school students meet systems that write, explain, flatter and keep them company. This paper argues that primary schooling cannot respond to that change through technological integration alone. The deeper question AI raises is not which tool to add, but how a school should spend students' time and what that time is for.

Drawing on current research on AI and education and on the Castlereagh Statement on AI in Australian education, the paper sets out a view of what primary education is for and derives from it a founding principle, ten principles and a working school design. The proposed PrimarAI day is the ARC Day: six sessions on ordinary NSW bell times, two each of Augment (AI-supported mastery practice in mathematics and English under a teacher's supervision), Reason (screen-free learning with a teacher, as a class) and Create (real-world learning organised through five Cs). Because the order rotates through the week, each part takes its turn at the first session of the day, when students learn most readily. The ARC Day is the latest of many versions of this school and the first that an existing school could trial with a single class. Digital literacy is taught as a subject from Kindergarten, and assessment is built into the fabric of everyday learning with the aid of purpose-built technology. The paper also sets out how adaptive AI coaching would work inside the Augment sessions, and shows two NSW Department of Education units re-cut for the ARC Day.

The model

These are the features that distinguish the proposed PrimarAI School from other NSW primary schools.

- **A proposed day in three equal parts.** Two sessions a day of AI-supported practice (Augment), two of screen-free learning with a teacher (Reason) and two given to uniquely human capacities (Create), with the order rotating so that each part takes its turn first thing.
- **Mastery progression.** Students work at their own edge within their Stage group. Nobody moves on until they have secured the skill, and nobody waits for the class to catch up. Each skill is taught through a six-move coaching cycle and is counted as secured only after a delayed check that a teacher samples.

- **Teachers at the centre.** Every adult leading learning is a NESA-accredited teacher. The AI is supervised, Socratic and purpose-built. It functions as a catalyst for thought, not a replacement.
- **Two Reason sessions a day.** Every day, with the screens off, teacher and class work together in the way current classrooms do at their best.
- **Digital literacy as a subject.** About an hour and a half a week from Kindergarten, in the two Reason sessions that open each Friday, with AI, its uses and its dangers as the subject rather than only the tool.
- **The five Cs in Create.** Communication, collaboration, creativity, character and connection, taught with real audiences, real budgets and real consequences.
- **Multi-age home groups.** About twelve students, with a teacher who stays with them across several years.
- **Six layers of assessment.** Mastery records, real performance, portfolios, self-assessment, narrative portraits and a graduation sign-off.
- **Fifteen graduation commitments.** Demonstrable things every leaver has done, with inclusive pathways for every student.
- **Privacy by charter.** Australian data residency, no model training on students' work, full parent access and a published annual audit.
- **No AI companions.** AI is a tool students study and direct, not a friend.
- **No points economy.** Progress is counted in syllabus outcomes secured. There is no currency, no prizes, no screen-time target and no camera on any student.
- **Open by default.** The whole design is published under a Creative Commons licence so others can test, copy and improve it.

How to read the claims in this paper

A paper like this mixes several kinds of statements, and readers deserve to know which is which. I use five categories throughout, and Part VI sets them out in full.

- **Established research** is a finding with a substantial and reasonably consistent evidence base, such as the effectiveness of well-run tutoring.
- **Emerging evidence** is a finding from the young and fast-moving research on AI in education, usually from a handful of studies with older students.

- **Reasoned extrapolation** applies established findings to a new situation, such as a primary school, where they have not been directly tested.
- **Design hypothesis** is an original PrimarAI proposal that needs to be tried and evaluated.
- **Forecast** is a prediction about technology, work or society. Forecasts are uncertain, and nothing essential in the design depends on any one of them coming true.

Preamble: Can you be mean to a chatbot?

My own interest in AI and students began in a Year 6 ethics class on a hot afternoon in 2023. I was a Primary Ethics volunteer then, teaching for thirty minutes once a week, and the topic that week was whether you can be mean to a chatbot. These were early days, when AI still looked like adults fussing over a new piece of software, yet the students took the question more seriously than I expected. The room went quiet, which is a strange sound in a primary classroom, and then everyone started talking over each other. Nobody else, it was obvious, was asking them anything like this.

I have thought about that afternoon a great deal since, because the question holds most of what this paper is about. It made me realise that the system was failing these students, and my intuition was that the question deserved to be taken seriously.

Why this paper exists

Three developments over the past two years have pushed this from a private project into something I think should be brought into the public realm.

The first is Alpha School, a private American network built on two hours a day of AI-supported academic work and an afternoon of life-skills workshops, now expanding fast and, in true Texan style, looking to take on the world. It was founded on the idea that AI makes forms of schooling possible that were previously unrealistic, and its founder has said that its central measure of success is that students are genuinely disappointed when the holidays arrive. It is expensive, selective, loudly marketed and contested, and the evidence for its results is much thinner than its publicity (De Bruyckere, 2026; Meyer, 2026). It has nevertheless done something brave. It has shown that a different shape for the school day is possible.

The second is Estonia's AI Leap, which since September 2025 has given about twenty thousand upper-secondary students and thousands of teachers access to AI tools adapted to ask questions rather than hand over answers, with teachers trained before their students were given access (Ministry of Education and Research, 2025; Thian, 2026). It is the most serious national attempt I know of to bring AI into schools on educational terms.

The third is the Castlereagh Statement, the product of a cross-sector summit hosted by the University of Sydney in October 2025 and published in 2026 as a call to action on Australian education in the age of AI (Castlereagh Summit, 2026). It sets out three goals, six principles

and three horizons of change, and it says plainly that "wait and see" is not an option (Castlereagh Summit, 2026, Trigger signals).

Meanwhile AI is already in NSW primary classrooms. Since Term 4 2025, NSWEduChat, the Department of Education's purpose-built chatbot, has been available to every public school student from Year 5 (NSW Government, 2025). The question is no longer whether AI enters Australian primary education, as it already has. We need to ask instead what shape the school takes around it.

These signals point to the same gap. We have national principles, international experiments and a working tool in our own classrooms, but no worked example of what a primary school designed for this moment would actually look like in NSW. This paper is an attempt to provide one. Primary is the right place to start, because it is where the early habits of using, questioning and doing without AI will be formed. Today's primary students are growing up as AI natives, and there are already calls for computing and AI to sit alongside reading and writing as basic skills ("Generation AI," 2026). How primary schooling shapes them therefore matters in a way it has not before.

The argument

The central claim of the paper is this.

> *Artificial intelligence is changing how knowledge is produced and found, and it is changing the conditions in which students learn to think, relate and take part in society. Primary schooling cannot be reconsidered through technological integration alone, because the real question AI raises is how a school should spend students' time and what that time is for. AI-supported practice can take on much of the individual, repetitive work of learning foundations. A school should use the time this releases to give students more of what only people and the real world can give them: teachers who know them, sustained reading, writing and talk, physical and creative work, and real responsibility in a real community.*

Put more simply, the aim is students who are confident with AI because they grew up directing it and questioning it, rather than dependent on it because they grew up next to it.

This is not an anti-AI document. If anything it is closer to the opposite. I think AI tutoring may turn out to be one of the most consequential developments in education for a long time, and I want students to grow up using these tools well as a matter of course. But I also think the future will reward students less for knowing how to use AI than for knowing what is

worth doing with it, and that the second kind of knowledge is built in ways no software can supply.

How the paper is organised

The paper is built around five questions, each of which creates the need for the next. Part I asks what has changed. Part II, the philosophical centre of the paper, asks what primary education is for, whatever the technological future. Part III asks what should change and what must remain, and answers with ten founding principles. Parts IV and V ask what kind of school follows, first by describing the proposed PrimarAI day, the ARC Day, and then by working through curriculum, teaching, assessment, inclusion, safeguarding and governance. Appendix A gives worked examples from two NSW units. Part VI asks how the model could be tested, and says in advance what results would count against it. Every feature of the school should follow from a clear view of what the school is for, so the design comes after the purpose, not before it.

Nobody has ten years of experience with AI in primary classrooms, because nobody has had ten years of AI in primary classrooms. That includes me. I am a NSW primary teacher who spent fifteen years in corporate facilities management before teaching, and who has been running AI literacy incursions in NSW primary schools since 2025.

Part I: Facing the future

1 What has changed

1.1 Knowledge, thinking and work

The easiest way to misread AI is to treat it as a faster search engine. A better description comes from philosophers of technology who argue that AI is an epistemic technology, one that changes what counts as knowledge, who counts as a knower and how knowledge gets made (Alvarado, 2023). In education, Chen (2025) goes further and describes generative AI as epistemic infrastructure: not a piece of equipment that teachers pick up and put down, but a layer underneath teaching, learning and assessment that quietly changes how all three work. One of his findings is a warning for schools. When he looked at AI tools for lesson planning and feedback, he found what he calls epistemic substitution, where the system completes the task efficiently but does not sustain the practices through which teachers build professional judgement. If that can happen to experienced adults, it can certainly happen to eight-year-olds. Selitskiy and Inoue (2025) go further again, arguing that education needs a paradigm shift towards the capabilities that keep humans ahead of AI, so that the task is less about integration than about discovering what remains.

Human beings have always offloaded thinking onto the world. We write lists, use calculators and ask people who know more than we do, and cognitive psychologists treat this offloading as normal and often sensible (Risko & Gilbert, 2016). What is new is the range of thinking that can now be handed over, and the fact that the thing receiving it talks back. Lodge et al. (2023) make the point well. Generative AI is not like a calculator, because it is interactive, generative and capable across almost any domain, so the relationship between learner and tool is closer to delegation or collaboration than to using an instrument. That is why the old answer, teach the skill first and then allow the tool, is harder to apply than it sounds.

Some of the pace of change can be measured. Researchers at METR have tracked the length of software tasks that AI agents can complete without help, and found that it has doubled roughly every seven months since 2019, possibly faster since 2024 (Kwa et al., 2026). The same researchers are careful to say that their tasks are cleaner than real work and that performance drops on messier problems. The rogue agents of the past few months have done nothing to allay concerns about these figures; if anything, they suggest the advance is exponential.

The effects on employment are less clear. Large language models touch a large share of occupational tasks (Eloundou et al., 2024), but early studies of actual labour markets have found small effects overall (Humlum & Vestergaard, 2025), concentrated losses for early-career workers in the most exposed jobs (Brynjolfsson et al., 2025) and modest economy-wide productivity gains in the near term (Acemoglu, 2025). Jobs and Skills Australia (2025) concludes that generative AI is more likely to augment Australian jobs than replace them.

My own experience sits awkwardly beside those averages. Outside teaching I work in the corporate world, and not long ago my employer announced, with some fanfare, an AI upskilling offer and a promise to retrain rather than let go. About two days later, around three hundred roles went to AI. My team barely talked about it, although everyone knew people who had lost their jobs. The mood was to keep your head down and hope for the best. One anecdote proves nothing about the economy, yet it does say something about how this change feels from the inside, and it raises a fair question for schools. Is keeping our heads down and hoping for the best the most we can do for our students?

From where I sit, the working world looks like a two-lane highway. There are people who use AI seriously and people who maintain the status quo, and the gap between them is widening. My own use of AI is serious rather than productivity theatre. It has turbo-charged my work in many places, and there are jobs I would not now choose to do without it. But I notice other things too. One colleague of mine hates AI and types very deliberately and slowly. His messages are short, sometimes aggressive, sometimes funny, occasionally rude, and I almost prefer them to the polished, bland word-wave that comes out of his power-user counterpart.

Beyond what can be measured are the forecasts. Leaders of the major AI companies have said publicly that systems exceeding human ability across most cognitive work could arrive within a few years (Amodei, 2024). The chat window itself may be transitional. In May 2026 Mira Murati's company Thinking Machines previewed interaction models that listen, watch and speak at the same time rather than waiting for a typed request (Wheatley, 2026). Today's chatbots at least make a student articulate a question, read an answer and decide what to ask next, and that friction is educationally useful. A system that watches, anticipates and interjects lowers the floor on how much thinking a person has to do before help arrives. By the time today's primary students are teenagers, the AI they live with may be ambient, in glasses, cars and walls. We may still be in AI's command-line era, when chat rewards verbal fluency and carefully crafted prompts, and a school that only prepares students for screen-

based AI risks preparing them for the past, the way a 1995 computing class taught DOS commands at the dawn of Windows 95.

The agent phase may prove just as temporary as the "prompt engineer" phase before it. Agents already have their own supervisors, other agents that monitor their work, learn from it and iterate, and each iteration moves the task further away from the human. If something like general AI does arrive, it may be less like a robot you can switch off than a colleague doing our work alongside us, and words like "tool" and "agent" may go to the same linguistic graveyard as "DVD." I am not sure what our role inside that will be.

I treat all of this as forecast, not fact. Some of it will be wrong, and claims that future systems will be sentient beings of some kind belong firmly in the uncertain column. The PrimarAI design does not depend on any of these predictions coming true. What it depends on is a simpler observation. Today, people running AI at scale still treat its output as a draft. Dan Hart, who was Head of AI at the NSW Department of Education while its EduChat tool was developed (Bashir, 2024) and whose team now runs hundreds of automated AI processes a day at a large Australian university, says it uses AI only where it is the right tool and estimates that its AI-generated first drafts are often about half wrong, though useful as a structure to work from (Bowen & Fleming, 2026). The comfort of saying that AI only produces a draft, or is only right half the time, will not last. The case for human learning has to be stronger than the claim that the machine still makes mistakes. Even if the machine could do almost everything, students would still need to know who they are, what is worth doing and how to bring something real into the world.

1.2 What AI is already doing to students' learning

In NSW the official version of AI in school is cautious and well designed. NSWEduChat runs on Department infrastructure, filters content and, in the government's words, "does not reveal full answers to students" (NSW Government, 2025). Hart has described the student rollout as among the largest deployments of generative AI to school students anywhere, with something like half a million students in scope (Bowen & Fleming, 2026). The unofficial version is the one students meet on a parent's phone.

The research on what that unofficial version does to learning is young, but one pattern keeps recurring. When AI does the work, performance goes up and learning does not necessarily follow. In a large field experiment in Turkish high schools, students given an unrestricted GPT-4 assistant solved far more practice problems, then did worse than the control group on an exam once the assistant was taken away. A version designed with teaching safeguards,

which gave hints rather than answers, largely removed the harm (Bastani et al., 2025). Fan et al. (2025) found that university students who wrote with ChatGPT produced better essays but learned no more than peers who had human or checklist support, and coined the phrase metacognitive laziness for the way learners stopped monitoring their own thinking. Stadler et al. (2024) found that students who researched a science question with an LLM experienced less mental effort than students using a search engine and produced shallower reasoning. A widely reported MIT preprint found weaker neural engagement and poorer recall among students who wrote essays with ChatGPT, although it involved a small sample and has not yet been peer reviewed (Kosmyna et al., 2025). Lehmann et al. (2024) found that the same tool helped when students used it to understand explanations and harmed learning when they used it to get answers.

The OECD's *Digital Education Outlook 2026* reaches the same conclusion at system level. Completing tasks with generative AI does not necessarily translate into learning, and purpose-built educational tools guided by clear pedagogical principles do better than general chatbots (OECD, 2026). Leslie Loble, one of Australia's most careful voices on digital equity, calls this the performance paradox, a short-term boost that is not retained when students are tested later, and warns that the risk is greatest for younger learners (University of Technology Sydney, 2026). Jason Lodge adds that unstructured AI use will "almost certainly widen existing equity divides" (Duggan, 2026).

The quality of this evidence deserves scrutiny. A meta-analysis of experimental studies found positive effects of ChatGPT on student performance (Deng et al., 2025), and a later audit of the same studies found that only a small minority described the treatment, the control condition and the learning measure clearly enough for the effects to be interpreted at all (Weidlich et al., 2025). Most of the research involves secondary or university students, and almost none of it involves children under twelve. In summary, the evidence pulls in both directions, but it consistently warns against confusing what a student can produce with AI for what a student has learned.

The technology journalist Joanna Stern, who spent a year using AI for almost everything (Stern, 2026), put the student version of this in human terms. Talking about the experience afterwards, she described a college student, Grace, explaining how her friends now study: AI does the research, the outline and the heavy thinking, and the students change a few words. "I now feel my brain dulling," Grace said, "and I don't like that" (Swisher & Galloway, 2026). The writing was never really the hard part. The thinking was being outsourced, and the student could feel it going quiet.

Two further features of consumer AI matter for students. The first is sycophancy. Language models trained on human approval tend to tell people what they want to hear (Sharma et al., 2023), and in April 2025 OpenAI withdrew an update to its main model after acknowledging it had become "overly flattering or agreeable" (OpenAI, 2025). When Stern tried an AI therapy app, her own therapist described it as more like a sycophant than a therapist (Stern, 2026). Frictionless agreement may feel pleasant, but growth does not happen that way.

The second is companionship. A widely shared *Harvard Business Review* analysis of what people say online about their AI use ranked therapy and companionship as the most common purpose in 2025 (Zao-Sanders, 2025). Large studies of actual conversations give a smaller figure. Affective conversations made up about three per cent of conversations on Claude, with companionship and roleplay under half a per cent (McCain et al., 2025), and self-expression was a small share of ChatGPT use (Chatterji et al., 2025). Among teenagers, though, the picture changes. In a 2025 national survey, nearly three in four American teenagers had tried an AI companion and about half used one regularly (Robb & Mann, 2025). Greg Brockman, OpenAI's president, has described a personal AI that knows your context, understands your goals and acts on your behalf as the consumer endgame of the technology (Noles, 2026). For an adult with a formed identity, that may be useful in real ways. For a young person who is still forming one, an always-available, always-agreeable presence is no longer really a tool. It has the shape of a relationship.

1.3 The shifting cultural ground

There is one more reason this matters now: the cultural ground is shifting. Stern's biggest takeaway from her year inside AI tools, as she has described it, was not that the technology got dramatically better but that public hostility to careless AI use got dramatically worse, particularly among young people (Swisher & Galloway, 2026). That backlash is the natural response of people who can feel the plates grinding under their feet. The people sounding the alarm about AI are not fools. They are the canaries, and the canaries are usually telling the truth. Rather than smashing the looms, the answer is to build the thing that makes the fear unnecessary.

There is a hunger for thoughtful design, and early signs of it are everywhere: in people buying handmade, reading printed books, going to live music and choosing the slow, deliberate human colleague over the polished AI-generated one, at least some of the time. That is good news for an Australian school launching now. A day in which four of the six sessions are spent with people rather than screens is in step with parents who want their children to be confident with technology and grounded in something more. The

conversation about how to do this in primary is one a lot of people have been waiting for, even if they have not quite known to ask for it in those words.

2 What schooling has achieved, and what it assumes

It is fashionable to say that everything about the industrial school is wrong. That overstates the case, and a proposal that starts from contempt for what exists usually ends up throwing away the wrong things.

Mass schooling produced near-universal literacy and numeracy, something almost no human society had achieved before. It built shared habits of attention, persistence and daily rhythm. It made school a civic place where students meet other students unlike themselves, and it turned teaching into a qualified, accredited and accountable profession. Any redesign worth taking seriously keeps four things from that inheritance: the expectation that every student, whatever their background, will reach strong foundations in literacy and numeracy; the discipline of a shared day and of finishing what you start; school as a common place; and the professional expectation that the adult in the room is a qualified teacher.

Some of the inheritance, though, is better described as administrative convenience than educational principle. The clearest example is age-grade progression, in which students born in the same year move through the same content at the same pace. It persists mostly because the alternatives are complicated to organise at scale, not because anyone seriously thinks year groups are the right unit for learning. Putting eight-year-olds in a room together because they happened to be born within the same twelve months is not obviously the best way to organise learning, and alternatives already exist, in the multi-age classes of Montessori, Steiner and small rural schools and in NESA's own Stage structure, which is already halfway there. The spread of achievement inside any one Australian year level is enormous. By Year 5 the middle sixty per cent of students span more than three and a half years of learning, and by Year 9 the top tenth are about eight years ahead of the bottom tenth (Goss et al., 2016). The NSW Curriculum Review recommended untimed syllabuses so that students could progress when ready rather than when the calendar says (Masters, 2020), and the Castlereagh Statement lists "schools organised by age cohorts" among the inherited structures that "were designed for a different era" (Castlereagh Summit, 2026, Principle 2).

The national results show what is at stake. In the national NAPLAN results for 2025, released by the Australian Curriculum, Assessment and Reporting Authority, two in three students reached the Strong or Exceeding proficiency levels and about one in ten needed additional support (The Educator, 2025). That is a system doing a lot right and still leaving a third of

students short of where they should be. The current timetable is built on an industrial compromise: one teacher, thirty students, one lesson, one pace. A whole-class lesson pitched at the middle of a room that spans several years of learning is a structural reason for that shortfall, not a failure of individual teachers.

Education has heard promises like this before. Sidney Pressey's teaching machines in the 1920s and B. F. Skinner's in the 1950s were sold on exactly the idea that every student could learn at their own pace (Watters, 2021). Film, radio, television and computers each arrived with predictions that they would transform classrooms, and each was absorbed into the existing routines of school (Cuban, 1986). Reich (2020) shows that new technologies tend to be domesticated by the institutions they enter and to benefit already-advantaged students most, and his recent podcast series with Jesse Dukes on generative AI in American schools documents how quickly AI homework help outran school policy, and how often schools have been left reacting to it (Reich & Dukes, 2025). I take that history seriously. It is the strongest argument I know for why this paper is about the architecture of the school day rather than about a product, because a tool dropped into an unchanged school gets absorbed by it.

3 Three responses, and what they leave open

3.1 Alpha School

Alpha School, founded in Austin, Texas, runs on a simple and bold idea. Students spend no more than two hours a day on AI-driven academic software and the rest of the day in workshops on life skills, entrepreneurship and public speaking. After about a dozen campuses it plans to reach roughly fifty in the 2026–27 school year, with tuition of US$40,000 to US$75,000 at most campuses (Meyer, 2026; Sands, 2026). Its own note to parents for the first session of 2026–27 counts 229 guides across 46 campuses (Alpha School, 2026). It employs guides rather than licensed teachers, and it reports that its students grow more than twice as fast as peers on nationally normed tests.

Those results have not been independently verified, and the school is selective, screening for engaged families and students within a typical range of ability and independence (De Bruyckere, 2026; Meyer, 2026). Unbound Academy, an Arizona public charter school using the model without selective admission, projected sixty per cent proficiency in mathematics in its first year and reported ten per cent, against a state average of thirty-four (Meyer, 2026). A school's first year proves little, but it may suggest that selecting successful students is much easier than making them.

I came to this work partly by listening to Alpha's founder talk on a podcast about putting student wellbeing first and being willing to admit mistakes, and I want to give the school its due. It has done the hardest thing in education, which is to run a different shape of day in public and let people argue about it. We have nothing remotely similar in Australia, where we wince at extreme entrepreneurialism, worry, rightly, about privacy and do not share an assessment culture built around college competition. Alpha is shaped by an American culture of testing, positional struggle, extreme entrepreneurial values and thin social support. None of those is a reason not to attempt an Australian version; they are reasons for it to look different. The Australian version cannot be Alpha, and it should not try.

3.1.1 How Alpha's two hours work

Because Augment borrows from Alpha's academic block, it is worth setting out how that block runs.

The morning opens with a short launch in which guides gather the students. Each student then opens an AI dashboard, called TimeBack, which shows what comes next in each core subject, chooses a first subject and starts without waiting for anyone else (Alpha School, n.d.-c). Work runs in short focus cycles, with a movement break (Pomodoro cycle) between the first hour and the second.

The dashboard is a front door to a set of separate practice applications. In 2025–26 these included Math Academy for mathematics from Grade 4, a fluency program for number facts in the early grades, and a reading sequence that moves from short AI-generated articles pitched close to the student's level, to personalised stories, to physical books. Guides do not teach during the block. Alpha describes them as coaches and motivators. In its youngest classes students track their own skills on check charts and sit a short test to earn each check.

Four design rules do most of the heavy lifting here. A student moves on only after reaching about 90 per cent on a concept, and importantly for motivation, lessons are tuned so that students answer correctly about 80 to 85 per cent of the time.

Progress is counted in points, with one point for each minute of productive work and a target of 120 a day, and whether a minute was productive is judged partly by software that watches how the student is working. Motivation here is engineered to a certain extent. Students earn a school currency, outdoor time and larger rewards for meeting goals, so there is clearly an extrinsic aspect to this.

Alpha has also iterated and updated, taking out unguided chatbots because the evidence so far suggests they get in the way of learning. That is the same reading of the evidence as Section 1.2 of this paper.

3.1.2 NextGen, Nova and Waypoint

Alpha's network now includes specialty schools that keep the two-hour academic block and change the rest of the day.

- NextGen Academy in Austin takes students from Grade 3 to Grade 8 and builds its afternoons around game-based workshops and gamified fitness, under the line "2 Hours to Learn, 4 to Triumph".
- Nova Academy, also in Austin, describes a lower-cost version in which students mentor each other and work with local organisations. This is a lot more similar to the PrimarAI School.
- Waypoint Academy, opening in Dripping Springs, is a nature-based school whose published day starts with journalling and a team huddle at 8:45, runs academics from 9:00 to a midday lunch, and spends the afternoon on survival skills, archery, strength training and a free block outdoors.

Therefore, the first is that Alpha itself treats the academic block as an engine and the rest of the day as a design choice, which is what the ARC Day does with Create. A school of bush, water, council chambers and market stalls is, in Alpha's terms, one more specialty, and Waypoint suggests there is an appetite for it. The second is that none of these schools publishes specific syllabus lesson examples: how a given syllabus outcome is taught, practised and checked across the day. Their pages give claims about speed and rank, but no lesson, no scope and sequence and no worked example. Sections 11 and 18 and Appendix A of this paper are an attempt to publish exactly that for a NSW school.

3.1.3 What PrimarAI takes, changes and leaves

Table 1

Features of the 2 Hour Learning schools and the PrimarAI position on each

Feature of the 2 Hour Learning schools	PrimarAI position	In the proposed ARC Day
Individual pacing with a mastery gate of about 90 per cent	Take, and tighten	Nine in ten on a fresh check, then a delayed check days later that a teacher samples (11.6)
Practice tuned to about 80 to 85 per cent success	Take	The tutor aims for about four correct in five (11.6)

Feature of the 2 Hour Learning schools	PrimarAI position	In the proposed ARC Day
Short focus cycles with movement breaks	Take	Two cycles of about fifteen minutes inside each 45-minute session (11.7)
A dashboard the student and family can see	Take, in plain language	A progress map against syllabus outcomes; families see everything (Section 23)
Themed, real-world afternoons	Take the idea	Create and the five Cs, anchored in NSW life (Section 14)
No ceiling on progress	Take, with a teacher's sign-off	Extension within the Stage first, then the next Stage (18.4)
Guides who motivate but do not teach	Change	NESA-accredited teachers who teach small groups during the block (11.8)
AI-generated lesson content	Change	A teacher-checked item bank mapped to the syllabus, and a way for students to flag the tutor as wrong (11.10)
Progress counted in minutes of productive work	Change	Progress counted in outcomes secured, never in minutes (11.9)
Software carries all academic teaching	Change	Reason: teacher-led, screen-free English, mathematics and science (Section 12)
Two hours in one block every morning	Change	Two sessions a day in a rotating order (Section 10.1)
Software that watches the student's attention, including by camera	Leave	No cameras and no attention tracking; the teacher in the room notices (Section 25)
School currency, prizes and cash for results	Leave	No points economy, and no food or money tied to targets (11.9)

3.2 Estonia's AI Leap

Estonia's AI Leap took a different path. A national programme launched in September 2025 gave about twenty thousand students in Years 10 and 11, in nearly all upper-secondary schools, access to AI tools adapted by University of Tartu researchers to work as Socratic tutors that prompt with questions rather than supply answers (Thian, 2026). Teachers were trained before students had access (Ministry of Education and Research, 2025). The government restricts student data to Europe and treats students' AI chats as private correspondence. Ivo Visak, who runs the foundation behind it, describes the rollout with disarming honesty as "like building a plane while it's flying" (Thian, 2026). Its evaluation is still under way.

AI Leap is integration done seriously, with pedagogy, sovereignty and teachers first. It is also a secondary-school programme that leaves the structure of the school day as it was. It asks how AI can improve the existing school, which is a good question but not the only one.

3.3 The Castlereagh Statement

The Castlereagh Statement is the closest thing Australia has to a cross-sector consensus on AI in education. Seventy contributors, from vice-chancellors and principals to researchers, students and public servants, made it in their personal capacities, and it describes itself as a green paper inviting endorsement (Castlereagh Summit, 2026). Its three goals are a shared definition of human value with measurement to match, coherent pathways from early childhood to lifelong learning, and confident, critical and creative engagement with AI for every Australian. Its six principles run from redefining the educated, future-ready Australian to placing technology in service of pedagogy and trust. Three of them speak directly to primary schools: redefining what a future-ready Australian is, building curriculum around capability rather than content, and making technology serve pedagogy and trust.

Its most useful idea is the distinction between three horizons: urgent stabilisation now, structural transition in the medium term and new foundations further out. At the far horizon for schools it proposes to "restructure schooling so students pursue pathways aligned with their interests," with families and communities taking shared responsibility for students' "critical judgement, digital discernment, and healthy relationships with technology from the earliest years" (Castlereagh Summit, 2026, Appendix B). It also says the 2023 Australian Framework for Generative AI in Schools is "fundamentally a risk-mitigation tool" rather than a framework for transformation (Appendix A). I agree.

The Statement does not single out primary schools. Its schools material is written mostly from a secondary vantage point, with ATAR-style transition criteria and career pathways in view. The near-horizon actions for primary are sensible and small, and the far horizon, where schooling is actually restructured, is left as an exercise for the reader. PrimarAI is my attempt to do that exercise in public. It should be read as one proposed, contestable interpretation of what the far horizon could look like in a NSW primary school, built now and at small scale so that it can be tested.

3.4 Other systems

Other countries are moving faster on the curriculum question than on the structural one. From the 2025–26 school year the United Arab Emirates introduced AI as a formal strand from kindergarten to Year 12 in government schools, inside an existing computing course,

with ethics making up a quarter of the content (Rasheed, 2025). Beijing now requires at least eight class hours of AI education a year in primary and middle schools, and Guangdong requires six hours a year in Grades 1 to 4 (Australian Government Department of Education, 2025). These are significant moves. They are also small amounts of time added to an unchanged day.

4 Why integration is not enough

4.1 More than a tool

Read the Castlereagh Statement carefully and a quiet assumption runs through it. AI is a technology, and the job of education is to use it well. Learners should "work with it to enhance rather than replace their own capabilities" (Castlereagh Summit, 2026, Principle 1). Assessment should help students "use AI to augment rather than replace their thinking" (Principle 3). Technology is to be placed in service of pedagogy. Those are sensible sentences, and I would sign most of them. My disagreement is about what they leave in place. Even at the far horizon the Statement imagines teachers, classrooms, curriculum and credentials with AI threaded carefully through them. Integrate the tool, train the teachers, and the framework gets to keep going more or less as it was.

That treats AI the way the 1990s treated the internet, or the 1950s treated television, as a new instrument slotted into an existing frame. It risks the mistake the generals of 1914 made, fighting an industrial war on the principles of a past age. The Statement itself acknowledges that our inherited structures "were designed for a different era" (Principle 2) and then proposes to reform them. But it is hard to renovate a building when the ground beneath it has already moved. AI did not so much break the industrial school as turn the lights on, so we could finally see the cracks. In my view, the cognitivist-behaviourist scaffolding our curriculum is built on, the credentialing logic our schools are organised around and the content-transmission model our timetables enforce are not being enhanced by AI. They are being made obsolete by it. In my own industry, the people who treat AI as something to adapt to are being overtaken by the people who treat it as a reason to redesign their business. I think the same will be true of schools.

Economic history suggests why integration alone may not be enough. When factories first adopted electric power, they swapped the central steam engine for a large electric motor and kept the old layout of shafts and belts, and productivity barely moved for decades. The gains came only when factories were redesigned around many small motors, with new floor plans, new workflows and new jobs (David, 1990). General-purpose technologies tend to pay

off through reorganisation rather than substitution (Bresnahan & Trajtenberg, 1995). If AI is a general-purpose technology for cognitive work, the important question for schools is how the day should be reorganised now that one of its most time-consuming parts, individually paced practice, can be done differently, rather than how the tool can be fitted into the existing day.

When in history has a general-purpose technology simply enhanced the method before it? The car did not enhance walking, the pocket calculator did not enhance mental arithmetic and email did not enhance the handwritten letter. Even so, the usual binary, enhance or replace, is not quite right either. Socrates wrote nothing down and, in Plato's telling, warned that the new technology of writing would hollow out memory and leave the appearance of wisdom in place of the real thing (Plato, ca. 370 B.C.E./1995). He was not entirely wrong, because every genuine shift takes something from us on the way to giving us something else. Calculators did not end the teaching of arithmetic, and used well they did not harm it (Ellington, 2003), but they changed what was worth practising. The pattern is reallocation. Something is handed to the new technology, and the time and attention it frees go somewhere else. Schools are not very good at deciding deliberately where that somewhere else should be, and this paper is an argument that we should decide it on purpose.

4.2 Responsible use is necessary, and not sufficient

The Statement returns often to the responsible and effective use of AI. I want students to use AI responsibly. A system built mainly on responsible use, though, asks the most of the people least equipped to provide it. Self-regulated learning, the ability to plan, monitor and evaluate your own learning, is a set of skills that students have to learn and practise over years (Zimmerman, 2002). There was never responsible use of social media at scale, and the AI market is unlikely to behave differently. The tools are clever, workarounds are easy and you can ask an AI to design the workaround for you. Policing it is like chasing a tiger, and anything built only on responsible use is built on a moving floor.

The alternative is not prohibition but design. If the environment makes the right behaviour the easy behaviour, with purpose-built tools that ask questions instead of handing over answers, time protected from AI by structure rather than by willpower, and explicit teaching about AI so that students understand what they are refusing and why, then responsibility becomes something students grow into, inside a school that carries the structure until they can carry it themselves. Castlereagh makes a related point when it says educators "should not carry this alone" (Castlereagh Summit, 2026, Appendix B).

Part II: What is primary education for?

Part I described what has changed. Before deciding what a school should do about it, we need a view of what primary education is for that does not depend on guessing the technological future correctly. If the purpose of a school is settled by predictions about jobs, then every new forecast reopens the question and the school chases the latest one. This part sets out a view of primary education sturdy enough to survive being wrong about AI.

5 Education for now, not only for later

For sixty years, schooling has run on human capital theory, the idea that education exists to prepare young people for work (Becker, 1964). For a long while that bargain was honest, because the future appeared legible. A young person in 1975 could look down the tunnel and make out the shape of the work at the far end, the trade, the office, the profession, and school was the train that got you there. Unfortunately, that tunnel has now gone dark. The OECD, whose economic framing of education has shaped school systems for decades, now says the human capital model "prepares young people for yesterday's jobs, through cognitive examinations," and argues that in the age of AI "education must strengthen human agency, human meaning and human security" (OECD, 2025a).

It is tempting to replace one prediction with another and to say that, because AI will do the routine cognitive work, students should learn taste, judgement and agency because the AI economy will pay for those. That is still human capital theory, just with a new list of skills, and it makes students' experiences valuable only for what they will be worth later. A student should learn an instrument for a thousand uniquely human reasons, not because musicians might be hard to automate. A student should swim in the harbour because the harbour is there and swimming in it is part of a good Sydney childhood. Martha Nussbaum makes the wider case that an education reduced to economic growth starves the capacities a democracy depends on, such as imagination, critical thought and the ability to see the world from someone else's position (Nussbaum, 2010).

The economic future still matters, and I do not want to pretend otherwise. Greg Brockman describes compute as the central resource of the coming economy and its allocation as one of the biggest questions society will face (Noles, 2026). In my own workplace, the people getting the most from AI are no longer only the technical experts. They are orchestrators, people who run several AI processes at once, one drafting, one checking, one building, and who stitch the results together with judgement. A future like that will reward people who use

AI well, at scale and with judgement, which is most parents' hope for their own children anyway. A student who can prompt, verify and direct AI, and who can also work with their hands, hold a difficult conversation and lead a project to completion, is exactly the kind of adult that economy will want around.

That, however, is a welcome consequence rather than the reason. The PrimarAI School is justified mainly by what it gives students now. The six hours of the school day are real education, not preparation for an education that happens somewhere else later.

6 Foundations, old and new

Today's students are still treated, broadly, as receptacles of "foundational" skills, and the dogged focus on those skills is meant to place them, supposedly, beyond the reach of grand cultural shifts. The reasoning sounds solid, and the word "foundation" sounds, by its nature, solid, the kind of thing one would be foolish to question.

Many of those foundational skills, though, were designed for an industrial and early-information economy. They assume that humans need to memorise, retrieve, calculate, transcribe and follow procedures faster than machines can. The foundations students will need concern working with information more than producing it: interpreting it, questioning it, directing it, navigating it emotionally, governing it ethically and transforming it creatively.

None of that makes the old foundations disposable; if anything, it makes them more important. Specialised knowledge is powerful precisely because it lets people think beyond their own experience (Young & Muller, 2013), and critical thinking is not a free-floating skill. It depends on knowing a great deal about the thing you are thinking about (Willingham, 2007). A student cannot judge whether an AI's account of the water cycle is right without knowing something about the water cycle, and cannot direct a machine creatively in a domain they know nothing about. Far from being the old curriculum that AI makes obsolete, literacy, numeracy and knowledge are the preconditions for interpreting, questioning and directing AI at all. The new foundations are built on top of the old ones, not instead of them.

There is a second kind of foundation that schools lay without naming it. Every tool a student spends hours with teaches them something about how the world works. A sycophantic chatbot teaches a student that the world agrees with them, and a feed teaches its own rhythms. School is one of the few places where students are deliberately brought into human practices of conversation, disagreement, turn-taking, care and civic life. A school that lets AI

become the main medium of a student's day hands part of that formation to commercial products, whatever their safeguards.

7 What a primary school should cultivate

The future will reward students less for knowing how to use AI than for knowing what is worth doing with it. Human-centred activities such as making, performing, collaborating, debating, cooking, building and enterprise teach students taste, judgement, initiative and endurance, the capacities that make AI useful rather than dominant. Table 2 sets out the capabilities the proposed school is designed to cultivate. I have kept the list from earlier versions of this paper and added the foundations that list took for granted. Each item matters whatever happens with AI, and each has a home in the school.

Table 2

Capabilities the PrimarAI School sets out to cultivate

Capability	What it means	Why it matters whatever the technology	Where it lives in the school
Literacy, numeracy and knowledge	Fluent reading and writing, number sense, and a growing store of knowledge about the world	Every other capability, including judging AI, depends on them	Augment and Reason sessions, inquiry questions
Taste and judgement	Knowing what is good and what is worth doing	Machines can generate options; someone has to choose	Reason, digital literacy, the five Cs in Create
Agency	Setting your own direction without being told	Students need to experience their actions mattering	Projects, portfolios, mastery choices
Self-knowledge	Knowing what you want, how you learn and what you will defend	Knowing how you learn is what makes learning last	Self-assessment, home group, narrative reports
Asking good questions	Framing problems, not only answering them	Often more valuable than producing answers	Augment inquiries, the Friday argument
Ethical reasoning under ambiguity	Deciding when AI should be in the loop and when it should not	Values are not delegated	Digital literacy, ethics circles, mock trials
Relational depth	Really being with other humans, including difficult ones	Students become themselves through others	Home groups, collaboration, older Australians

Capability	What it means	Why it matters whatever the technology	Where it lives in the school
Creativity	Making something new that is worth making	It is a human good, and possible in every field and for every student	Music, making, theatre, writing
Embodied capability	Physical skill, manual craft, presence in a room	Bodies are not optional, and skill is satisfying	Physical challenges, making, outdoor learning
Civic participation	Knowing how a community works and how to act in it	Democracy is learnt by doing it	Council visits, letters, community projects
Synthesis	Pulling threads together across the domains AI handles separately	Real problems do not arrive sorted by subject	Create projects, the loop between Augment, Reason and Create
Adaptive capability	Reading a situation, learning what is needed quickly, applying judgement, working with others and trying again when it goes wrong	No one can confidently name the skills these students will need at twenty-five	Character strand, long projects, real consequences

The last row is the one that ties the others together. What this school is aiming at, in the end, is graduates with deep adaptive capability. That does not mean students who have completed a checklist of skills, because the skills they will need at twenty-five are not the skills any of us can confidently name in 2026.

Agency deserves a further word, because it is the capability most often named by the people who build AI. Brockman names agency as the skill that matters most, describing it as knowing what you want, being able to explain it clearly and being able to judge whether the result is good (Noles, 2026). He is right about the capability and wrong in assuming that it forms by itself, because it does not. Agency grows when students have real autonomy, real competence and real relationships, the three needs at the heart of self-determination theory (Ryan & Deci, 2000). Vision, taste and the ability to know what is worth building are formed through years of touching real things, making real things and watching real things fail, and they cannot be acquired later from a chatbot, no matter how patient or capable that chatbot is. A childhood spent only in front of screens tends to produce students who can prompt without being able to taste, and the gap shows up later.

8 The value of friction

Students become themselves by bumping up against the world. Everyone in 2026 is being sold frictionless learning, frictionless friendship and frictionless answers. A primary school's

job is to choose the right kinds of friction, deliberately, and to protect them from being smoothed away by tools that, however helpful in other contexts, would replace the rub a student actually needs. That claim needs care, though, because not all difficulty is educational.

The research on learning distinguishes difficulty that helps from difficulty that hurts. Desirable difficulties, such as retrieving something from memory rather than rereading it, spacing practice over time and mixing problem types, feel slower and produce more durable learning (Bjork & Bjork, 2020). Productive failure, in which students wrestle with a problem before being taught the method, can deepen conceptual understanding when it is carefully designed and followed by instruction (Kapur, 2016; Sinha & Kapur, 2021). On the other hand, novices left to discover foundational knowledge with minimal guidance usually learn less, because their working memory is overwhelmed (Kirschner et al., 2006). Critics respond that well-scaffolded inquiry is not minimal guidance at all (Hmelo-Silver et al., 2007). That debate is unresolved, and I do not want to pretend otherwise. The sensible reading is that structured, explicit support works best when students are learning new foundations, and that productive struggle earns its place once students have enough knowledge to struggle with.

There is also friction that should be removed. When Alpha students are asked what they like about their AI tutor, the one that knows them well and encourages them tirelessly, the answer I keep hearing is that it *does not judge them*. This is consistently underestimated as a strength of AI tutoring, and whatever else can be said about AI tutors, students' experience of a good one is often overwhelmingly positive. Students who see asking for help as a threat to how others see them often avoid asking, and the students who most need help are the most likely to avoid it (Ryan et al., 2001). An AI tutor has no ego and no hidden agenda; it does not move on because the lesson plan says to, and it can sit with a student on the same fraction problem for forty-five minutes without sighing. That removes a kind of friction that was never educational, only embarrassing.

This gives a clear design rule. In the proposed design, Augment is built to remove unproductive friction: the mismatch between a student's level and the class pace, the embarrassment of not knowing, the waiting. Reason and Create are built to add productive friction: real audiences, real budgets, teammates who disagree, drafts that come back covered in comments, and consequences when something fails. We want students to feel safe to be wrong and still to feel the expectation to keep going. Frictionless interaction does not produce growth, however pleasant it feels in the moment, and sycophantic AI is the most

frictionless interaction of all. That is the best reason I know to keep it away from the parts of the day where students are becoming who they are.

9 What must remain human

If AI can do more and more of the explaining, practising and checking, what is left that must stay human? I think five things, and they set the limits on everything that follows.

The first is **deciding what is worth learning and why**. Curriculum is a set of choices about what a community wants its young people to know and become. Those choices are political and ethical, and they belong to people.

The second is **knowing a student over time**. There are things only a person in a room can do for a student. They can read a face, notice that a quiet student has gone quieter this week and make a class feel like a family. These are the things that make a school a school, and they have been getting squeezed out of teaching for decades by the workload of content delivery and the administration around it. The research supports this. Close, warm relationships between teachers and students are consistently associated with engagement and achievement (Roorda et al., 2011), the quality of those relationships in the early years predicts outcomes years later (Hamre & Pianta, 2001) and being taught again by a teacher who already knows you has measurable benefits (Hill & Jones, 2018).

The third is **holding a standard with care**. A teacher can hold a high standard when a student wants to give up. The effects of feedback vary enormously with how it is given and received (Hattie & Timperley, 2007; Wisniewski et al., 2020), and critical feedback works best when a student trusts that it comes from someone who holds high standards and believes they can meet them (Yeager et al., 2014). A teacher who reads a student's writing line by line, and whom the student cannot bluff, is doing something different from a system that scores the writing, even if the system's comments are good.

The fourth is **growing up among other people**. The work of childhood is figuring out who you are by bumping against other humans who do not always agree with you. That is why the chatbot question in the Preamble matters, and why the moment a nine-year-old stands up in front of adults at a council meeting and has to say what she thinks matters. An AI tutor, however patient, can rehearse a student for those moments, but it cannot stand in for them.

The fifth is **the final word on any judgement about a student**. Decisions about a student's progress, support, wellbeing or future are made by teachers and families. AI can inform those decisions, but it can never make them.

From these follows the design principle for this school, straightforward enough to say in one line. AI is a tool, not a friend. That is a rule for students, grounded in what we know about child development and online safety (Kurian, 2025; Robb & Mann, 2025), and it is not a claim about what AI is or might become. Whether future systems could have something like experience is an open question, and the digital literacy strand invites students to argue about it. Reason and Create are where students build the human relationships that actually make up a self, with teachers who know them over years, a multi-age home group they grow up with, sustained friendships with older Australians and classmates working on something difficult together. AI replaces none of these, and in fact makes more room for them, because it can now *take on parts of the day that school used to spend six hours on*.

Part III: The principles of an AI-native school

Part II set out what primary education is for. Part III turns that view into a founding principle and ten principles that follow from it.

The founding principle can be said in one sentence. Anything an AI can teach a student well, the AI should help teach; anything it cannot, a teacher should; and we should make far more time for the second than we currently do. Even where AI is good at practice, a teacher still teaches the foundations first and decides when the machine is helping and when it is not.

Two consequences follow. First, the creative parts of the curriculum should be brought forward and the process and content parts given less of the day. AI is very good at process and content, yet it is weaker at original creative direction, and a student who can direct it creatively is going to be a different kind of citizen from one who can only follow it. Second, the whole curriculum needs to be reframed as learning to learn. Metacognition is what lasts, which is to say knowing what you know, knowing how you learn, knowing when you are being deceived and knowing when to push and when to pause.

Each of the ten principles below gives its justification, the research it rests on and what it means in practice. The principles come before the design on purpose. Every major feature of the school in Parts IV and V should be traceable to one or more of them, and a feature that cannot be traced should be cut.

Principle 1: Purpose comes before technology

Why. A new timetable is only an outward arrangement. It is worth something only as the expression of a changed idea of what students' time is for. Without that, it is just a different way of running the same school.

Evidence. A century of educational technology shows that tools introduced without a change in purpose and structure are absorbed into existing routines (Cuban, 1986; Reich, 2020).

In practice. The school's design is derived from the purposes in Part II, and its use of AI is justified case by case against them. Where AI does not serve a clear educational purpose, it is not used.

Principle 2: Every student learns at their own edge

Why. A school that takes each student seriously cannot pitch every lesson at the middle of a room that spans several years of learning.

Evidence. Mastery learning, in which students work towards a defined standard with feedback and corrective teaching before moving on, has a long and positive research record (Bloom, 1968; Kulik et al., 1990), though effects are smaller on standardised tests than on teacher-made ones (Slavin, 1987). Achievement within a single year level in Australia spans several years (Goss et al., 2016).

In practice. In Augment sessions students progress through NSW syllabus outcomes at their own pace within their Stage group, with mastery checked before they move on.

Principle 3: Teachers lead, and AI stays in the loop

Why. The teacher's judgement, relationships and authority are the parts of education that must remain human (Section 9). AI should extend a teacher's reach, never replace their responsibility.

Evidence. The strongest AI tutoring results so far come from systems with humans in the loop. Tutors given an AI assistant in a large American trial helped more students master topics, with the biggest gains for the lowest-rated tutors (Wang et al., 2025). In UK classrooms, an AI tutor whose messages were checked by expert tutors performed as well as the tutors alone, and tutors approved about three-quarters of its drafts with little or no editing (Google LearnLM Team & Eedi, 2025). The Castlereagh Statement's fifth principle insists on teachers' agency over when and how AI is used.

In practice. Every adult leading learning is a NESA-accredited teacher, and the word used for them is always teacher. The AI flags, suggests and drafts. Teachers decide.

Principle 4: Foundations are taught explicitly, and by people where it matters most

Why. Literacy and numeracy are preconditions of judgement (Section 6). They are too important to leave to chance or to software alone.

Evidence. Learning to read depends on systematic, explicit phonics teaching followed by extensive reading for meaning (Castles et al., 2018). Explicit, well-sequenced instruction helps novices (Rosenshine, 2012). Not every digital phonics programme works: a rigorous English trial of the phonics game GraphoGame Rime found no improvement in struggling Year 2 readers' reading (Education Endowment Foundation, 2018). Children understand stories read from paper better than the same stories read from screens, and adult shared reading helps more than digital enhancements children use alone (Furenes et al., 2021).

In practice. Teachers teach phonics and early number explicitly. Adaptive software provides practice, not first teaching, in Kindergarten to Year 2. Books are on paper.

Principle 5: Independent thinking comes before assisted thinking

Why. A student cannot direct or check an AI in a domain where they cannot think unaided. The aim is confidence with AI, not dependence on it.

Evidence. Unrestricted AI help raises performance on practice tasks and can lower learning once the help is removed, while tools that prompt rather than answer reduce the harm (Bastani et al., 2025; Fan et al., 2025; Lehmann et al., 2024). Offloading is sensible only when the offloaded skill is not the one being learnt (Risko & Gilbert, 2016).

In practice. The tutor asks questions rather than giving answers. Large parts of the day are protected from AI by the timetable. Students are taught explicitly when to close the lid and do the thing themselves.

Principle 6: AI is studied, not only used

Why. Students cannot question what they do not understand. They will pick up habits around AI whether or not school teaches them, so those habits should be formed deliberately rather than by accident.

Evidence. AI literacy frameworks for school students now exist from UNESCO (Miao & Shiohira, 2024), the European Commission and OECD (2025) and the AI4K12 initiative (Touretzky et al., 2019), building on research about what AI literacy involves (Long & Magerko, 2020). Castlereagh's third goal calls for developmentally appropriate learning about AI for every Australian.

In practice. Digital literacy is a subject with its own time each week and its own Kindergarten to Year 6 progression, taught by teachers, with AI as the object of study.

Principle 7: Relationships and real experience are protected by design

Why. The work of childhood is figuring out who you are by bumping against other humans who do not always agree with you. That needs real others and real stakes, and it needs time that cannot quietly be traded away for something more measurable.

Evidence. Teacher–student relationships matter for engagement and achievement (Roorda et al., 2011). Experiences in nature are linked to learning, attention and wellbeing (Kuo et al., 2019). Well-designed project-based learning can improve achievement in primary science (Krajcik et al., 2023).

In practice. Two screen-free Reason sessions and two Create sessions every day, so that four of the day's six sessions are spent with people rather than screens, multi-age home groups that stay together and a teacher who stays with them.

Principle 8: No student is designed out

Why. A school that claims to take each student seriously must work for a diverse student cohort.

Evidence. Australian law requires reasonable adjustments so that students with disability can participate on the same basis as their peers (*Disability Standards for Education 2005*). Universal design for learning offers a framework for building flexibility in from the start (CAST, 2024). Edtech tends to widen gaps unless designed against it (Reich, 2020).

In practice. Every requirement in this paper, including the graduation commitments, has an inclusive pathway. A scholarship pathway exists from day one.

Principle 9: Trust is earned through privacy, safeguarding and accountability

Why. Parents asked to send their children into an AI-rich school are owed protections by default, not by request.

Evidence. UNESCO's guidance on generative AI proposes thirteen as a minimum age for independent use of AI tools in classrooms (Miao & Holmes, 2023). Children are especially likely to misread chatbots as caring (Kurian, 2025). Australian privacy law is being strengthened for children, including a Children's Online Privacy Code due by December 2026 (Office of the Australian Information Commissioner [OAIC], 2026). Estonia shows that data sovereignty can be made a condition of national partnerships (Thian, 2026).

In practice. Charter commitments on data, age-appropriate design that limits open-ended conversation for younger students, teacher-first rollout and independent audit (Section 25).

Principle 10: Students learn by taking part in community life

Why. Students learn best by friction against the world, and the world they will inhabit is outside the classroom. A lunch that Year 5 cooked is not a break from the programme but part of it, and a letter to the council that gets a reply teaches more about civic life than a worksheet about councils.

Evidence. Project-based learning shows positive average effects on achievement, with wide variation by design (Chen & Yang, 2019), and a well-designed project-based science

programme improved primary students' achievement in a randomised trial (Krajcik et al., 2023).

In practice. Real audiences, real budgets, real customers, real councils and real older Australians, every week.

Part IV: The PrimarAI day

10 The ARC Day

The proposed PrimarAI day is the ARC Day. It has three main parts, Augment, Reason and Create, and it runs on ordinary NSW bell times, with six sessions between 9:00 and 3:00 and two sessions of each part every day. Because the bell times are ordinary, the strategy can be retrofitted into current primary school structures.

Augment is AI-supported practice in English, mathematics and HSIE, content-rich subjects that benefit from both teacher-led sessions and finely tuned AI coaching. That practice sits at each student's own edge, with teachers in the room.

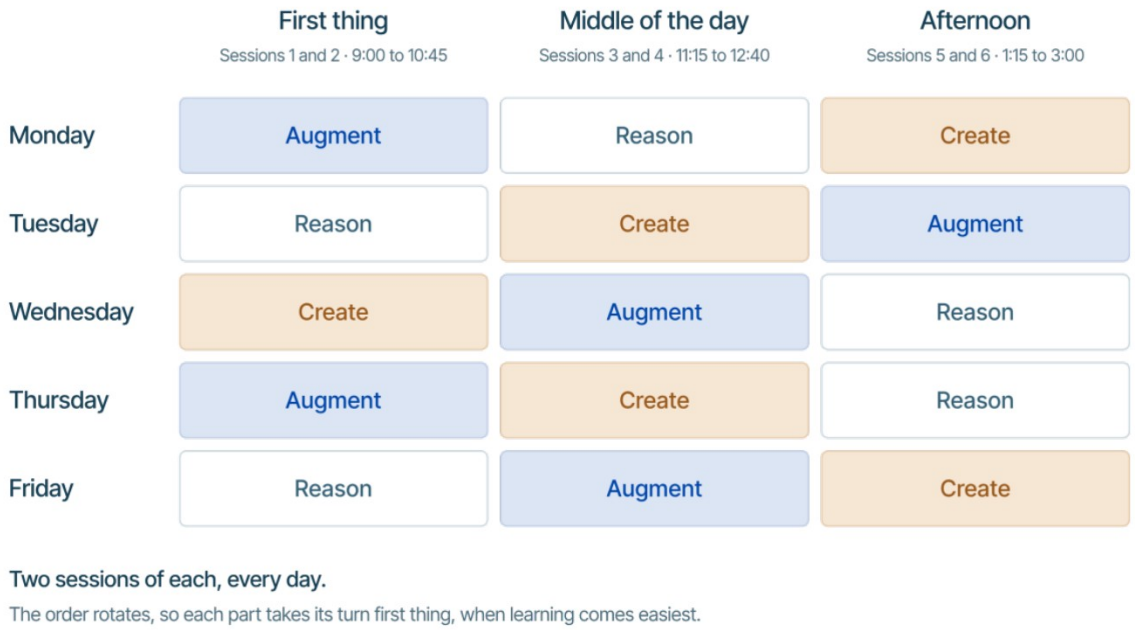
Reason is learning with a teacher, as a class and with the screens off. It carries English, mathematics and every science and technology lesson, because the practical nature of science in particular pays off most in the hands of a teacher.

Create is the human factor: the creative arts, PDHPE and sport, team building, making and enterprise, out in the world as often as in the building. In the proposed timetable the sessions run in pairs, before recess, before lunch and after lunch, so that students have time to settle into a piece of work, and the order of the pairs rotates through the week so that each part takes its turn first thing.

Alpha proved that a different shape for the school day was possible. This is a different shape again, and every part of it comes from the principles in Part III. Augment applies Principle 2 (every student at their own edge) under Principle 3 (teachers lead). Reason applies Principles 4 and 5 (foundations taught by people, and independent thinking first). Create applies Principles 7 and 10 (real relationships and community life). Friday's digital literacy session applies Principle 6.

Figure 1

The proposed ARC Day rotation (example week)



Note. Each row is a school day and each block is a pair of sessions. In the example week, Augment goes first on Monday and Thursday, Reason on Tuesday and Friday, and Create on Wednesday.

Table 3

The proposed daily timetable (example)

Time	Block	What happens
8:30–9:00	Arrival	Students arrive, most walking themselves in by Year 2
9:00–9:45	Session 1	Roll in home groups, then the first pair of sessions
9:45–10:00	Crunch and sip	Fruit, water and a movement break
10:00–10:45	Session 2	First pair continues
10:45–11:15	Recess	Unstructured time outside, with no phones
11:15–12:00	Session 3	Middle pair
12:00–12:40	Session 4	Middle pair continues
12:40–1:15	Lunch	Eaten outside; cooked by a student team once a fortnight
1:15–2:00	Session 5	Afternoon pair
2:00–3:00	Session 6	Afternoon pair continues
3:00	Home time	Extended care available until 5:30

Table 4

The proposed ARC week (example)

Session	Monday	Tuesday	Wednesday	Thursday	Friday
Goes first	**Augment**	**Reason**	**Create**	**Augment**	**Reason**
Session 1 9:00–9:45	**Augment** English	**Reason** English	**Create** Music	**Augment** English	**Reason** Digital literacy
	Crunch and sip, 9:45–10:00				
Session 2 10:00–10:45	**Augment** Mathematics	**Reason** Mathematics	**Create** Health and wellbeing	**Augment** Mathematics	**Reason** The Friday argument
	Recess, 10:45–11:15				
Session 3 11:15–12:00	**Reason** English	**Create** PE	**Augment** English	**Create** Visual arts	**Augment** Mathematics
Session 4 12:00–12:40	**Reason** Mathematics	**Create** Drama and dance	**Augment** Mathematics	**Create** Team games with buddies	**Augment** History and geography
	Lunch, 12:40–1:15				
Session 5 1:15–2:00	**Create** Assembly, then team building	**Augment** English	**Reason** English	**Reason** Mathematics outside	**Create** Community and showcase
Session 6 2:00–3:00	**Create** Making and enterprise	**Augment** History and geography	**Reason** Science and technology	**Reason** English, library and long writing	**Create** Sport

Note. The coloured label in each cell shows the part of the day; shaded rows are breaks. One Tuesday a fortnight the whole school goes off-site for the day instead. Table 8 gives the minutes and share of the week for each learning area in this example.

The three parts are meant to work as one system, not three programmes that share a building. Augment gives teachers detailed information about each student, and Reason is where teachers come to know the student behind the data. Inquiry sessions in Augment send questions into Reason and Create, and Reason and Create send measurements, writing and problems back into Augment. Section 15 describes that loop in more detail.

10.1 Why the first session matters

Every teacher knows that the first session of the day is where the real learning happens. Students arrive rested and relatively focused, the day is fresh and new ideas have a chance to stick, and by the last session the same lesson is much harder going. Experienced primary

teachers protect their mornings for this reason, and conventional timetables usually give them to English and mathematics as a matter of course.

The research supports what teachers see. Using records for nearly two million students in Los Angeles, Pope (2016) found that students learned more in morning classes than in the same classes later in the day, with the largest effects in mathematics. Sievertsen et al. (2016) analysed Danish national tests taken by children aged eight to fifteen and found that scores fell with each hour later in the day that a test was sat, and that a break of twenty to thirty minutes lifted them again. Neither study tested a rotating timetable, and most of Pope's students were older than primary age, so applying these findings to the ARC Day is reasoned extrapolation. The two studies are, though, consistent with each other and with what teachers report.

In the proposed ARC Day each part takes its turn. In the example week, Augment goes first on Monday and Thursday, Reason on Tuesday and Friday, and Create on Wednesday, so that across the week each part has between seven and a half and eight hours; a three-week cycle would share the first sessions exactly.

The sessions are kept in pairs for a related reason. Sessions that change subject every forty minutes leave little time for a student to settle into hard work, and the fatigue evidence suggests that breaks restore attention better than a change of activity alone (Sievertsen et al., 2016). Each pair belongs to one part of the day, crunch and sip falls inside the first pair, and recess and lunch fall between the pairs.

10.2 How the design got here

The ARC Day is the latest of many versions of this school. Each version answered a weakness in the one before, and the sequence deserves to be set out openly.

The first edition of this plan, in June 2026, borrowed the shape of Alpha's day almost whole: a two-hour morning of AI-supported practice and a four-hour afternoon of projects and life skills. The second edition, in September 2026, cut the afternoon to three hours and put a screen-free human hour between the two, because the evidence on writing, reading on paper and teacher feedback made it clear that the meaning-making half of English could not be left to software or folded into projects (Furenes et al., 2021; Graham et al., 2012). It also made digital literacy a subject with its own hours.

Working the two, one, three day through in detail exposed three problems. The first was the one described above: the best hours of every day went to the same part of the programme.

The second was mathematics. Measured against NESA's time allocation advice, the September timetable gave mathematics 20 per cent of the week only if two hours of it were deliberately programmed into afternoon projects, which is the kind of promise that is easy to make and hard to keep (Section 18). The third was that the day only fitted a school built for it. An 8:45 morning circle, a two-hour block before recess, a single hour before lunch and a home-group half hour after 3:00 do not map onto the bell times of an ordinary NSW primary school, so the model could only ever have been tried in a new one.

The ARC Day keeps the founding principle, the five Cs, the six layers of assessment, the graduation commitments, the privacy charter, the home groups and digital literacy as a subject, all of which the earlier versions got right. What it changes is how the time is divided. Each part gets two sessions a day, the order rotates, mathematics and English have their own sessions in both Augment and Reason, and the whole day fits the bell times NSW primary schools already use.

10.3 Starting inside an existing school

The most important consequence of the change is that the ARC Day does not need a new school. An ordinary NSW primary day runs from about 9:00 to 3:00, with crunch and sip, recess and lunch, and the proposed ARC Day uses those same times (Table 3). Nothing about the building or the length of the day has to change for one class to try it, and the design is meant to be adopted in parts as well as whole.

What makes this possible now is the technology. Individually paced practice used to depend either on a teacher preparing work at several levels at once or on a long, uninterrupted block in which software could carry each student forward. A purpose-built AI tutor that knows where each student sits against the syllabus can pick up where the student left off in any forty-five-minute session, so mastery practice no longer needs a two-hour block of its own. Alpha's own block is run in short focus cycles with a break between the first hour and the second (Alpha School, n.d.-c), so two 45-minute sessions are closer to its practice than they first appear. That fits the tutoring evidence, which finds the largest effects for programmes run during the school day and at least three times a week (Nickow et al., 2024). Two Augment sessions a day meet both conditions.

For an existing school the requirements are modest and specific. Augment needs a tutor that meets the privacy and design commitments in Section 25, devices for one class and a term for the teachers to learn the system before the students do. Reason is close to what good primary teachers already do, taught with the screens off. Create draws on time the school

already gives to the creative arts, PDHPE, sport, assembly and specialist lessons, and special religious education and special education in ethics can sit inside a Create session. The first test of the whole day, set out in Section 30, is therefore one class in one existing school for one term.

11 Augment

The NSW Mathematics K–10 and English K–10 syllabuses are the spine of Augment (NSW Education Standards Authority [NESA], 2022a, 2022b), and history and geography inquiry sessions make up the rest (Section 11.3). An AI tutor, purpose-built, hosted in Australia and compliant with the Australian Privacy Principles, knows which Stage outcomes each student has secured and which are next, and sets practice at the student's actual edge rather than the class median. Over time it learns how that particular student thinks: what they are curious about, where they got stuck yesterday, why fractions clicked as pizza slices but not on a number line. In most classrooms a student receives individual attention only briefly, and often only when struggling loudly enough to be noticed. My own daughter is lucky to get five minutes of it in a month. The proposed Augment sessions are designed to change that equation, and the people building the models are betting the same way. In May 2026 Anthropic and the Gates Foundation committed US$200 million over four years to AI in health, education and economic mobility, including evidence-based tutoring for school students (Anthropic, 2026). Sections 11.5 to 11.12 set out how that tutor would work, move by move, and Appendix A applies it to two NSW units.

Mastery means that no student moves on until they have secured the skill, and no student is held back because the rest of the class has not caught up. Mastery is checked, not assumed. The tutor tracks practice, but a skill is recorded as secured only when the student can show it on a fresh task days later, and teachers sample those checks.

11.1 What a proposed Augment session looks like in Year 4

In the proposed design, a Year 4 student sits at a workstation with headphones on. An Augment session lasts forty to sixty minutes, most often forty-five. It begins with five minutes of retrieval practice on something learnt earlier, which strengthens memory and lets the system calibrate (Bjork & Bjork, 2020). Then comes about thirty minutes of new learning at the student's edge, in two cycles of about fifteen minutes with a minute to stand and stretch between them, where the tutor uses Socratic prompts in the Estonian style rather than direct answers. If the student is stuck for more than two minutes, the system flags it on the teacher's screen and the teacher comes over to talk it through. The session ends with a short

check. Between the two sessions of a pair there is either crunch and sip or a short movement break, and the subject changes, usually from English to mathematics.

Teachers are not supervising from a desk. In the Stage 2 room about 26 students work with two teachers, and a learning support teacher moves between the rooms. The teachers move between workstations, pull small groups aside for explicit teaching when several students hit the same wall, run short conferences with students who have just finished a unit, and watch for the student who is gaming the system. There is no screen-streaming to anyone outside the school, no advertising and no data leaving Australia.

11.2 Kindergarten to Year 2 is different

Two sessions a day at a screen are not appropriate for a five-year-old, and the design does not ask for it. UNESCO proposes thirteen as a minimum age for independent use of generative AI in classrooms (Miao & Holmes, 2023), and NSWEduChat starts at Year 5. In Kindergarten to Year 2, Augment is mostly taught by teachers in small groups: systematic synthetic phonics, early number, handwriting and oral language, with concrete materials and plenty of talk. Adaptive practice software is used in short sessions, about forty-five minutes across the two Augment sessions in total, for the kind of rapid, individual practice young students benefit from: phoneme–grapheme correspondence, blending, number facts. It is closer to a well-designed practice game than to a chatbot, and it does not hold open-ended conversations. Alpha's own experience points the same way. Its ideal ratio in its youngest classes is one guide to six or eight students, and it tells parents that what motivates a five-year-old most is time with their guide. Appendix A.1 shows what a teacher-led Augment session looks like for a Stage 1 English unit.

Conversational features are introduced gradually in Stage 2, under close supervision, and used more fully in Stage 3. That progression is itself a design hypothesis to be evaluated (Section 30).

11.3 Inquiry sessions

Augment is mostly mathematics and English, but not only. Twice a week in the example timetable an Augment session is given to history and geography as an inquiry, where the tutor works through HSIE content at the student's level, and its job there is different. Rather than delivering content, it helps the student generate questions they cannot answer at a workstation. Why did the compost in bed three run hotter than bed one? What did the council actually decide about the crossing, and who voted against it? What will the tide be

doing during the harbour swim on Thursday? Those questions get written down and go to Reason and Create, where the answers live.

11.4 What Augment claims, and what it does not

Augment gives a Stage 2 or Stage 3 student about six hours a week of individually paced practice in mathematics and English, plus about an hour and forty minutes of history and geography inquiry. Earlier versions of this paper claimed this would produce more learning than a full week in a conventional classroom. That may turn out to be true, but it is a hypothesis, not a finding. The best evidence on tutoring (Nickow et al., 2024) and on AI tutors with humans in the loop (Google LearnLM Team & Eedi, 2025; Wang et al., 2025) makes it plausible, while the evidence that unrestricted AI help can harm learning (Bastani et al., 2025) makes it conditional on design. There is also almost no direct evidence for primary-age students using conversational AI tutors. Whether two well-designed sessions a day are enough is exactly what a pilot must test. The detail that follows does not change that, and it makes the hypothesis specific enough to build and to test.

11.5 What the tutor knows - the syllabus map

Adaptive AI coaching starts from a specific mapped pedagogy iterated by trained teachers. Before any student logs in, teachers separate each NSW syllabus outcome into its content points, which NESA already provides, and then into skill steps small enough to teach, practise and check in one sitting.

The Stage 2 outcome MA2-AR-01, for example, includes a content point on selecting strategies flexibly for addition and subtraction of numbers up to three digits. On the map that becomes six steps: use compensation to add, use compensation to subtract, use levelling, use constant difference, choose between them for a given problem, and explain the choice. Each step is linked to the steps it depends on, so a student who cannot yet bridge to the next ten is not asked to level 28 + 35.

For each step the map holds three aspects, the syllabus reference it serves, the steps that come before it, and how it is checked. Everything the AI tutor does is a move on this map, which is also what makes the tutor accountable to the syllabus. A teacher, a parent or a NESA inspector can pick any step and see which outcome it belongs to.

Where the students academic level is the most important starting point. In a student's first week the tutor runs short diagnostic sets, ten minutes at a time across several sessions, instead one long test which is usual in usual schools, and the teacher checks the result against the student's work samples and the school's existing assessment data. Each student

then starts at their first unsecured step in each strand, which may be below, at or above their Year level. In a Year 4 room at the PrimarAI School, that range will be several years wide, so it takes the scaffolding theory seriously. I have taught a composite Stage 2/3 class and can see in real time how the Stage 2 student benefit from the more advanced shared classes. So much so I had to literally slow the pace as they would have been bored the next year.

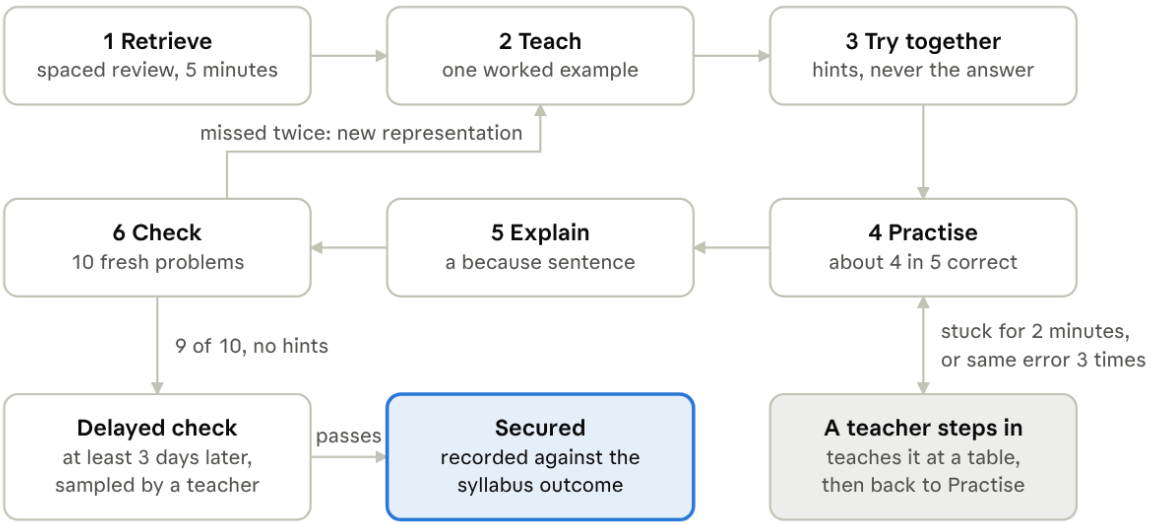
The tutor keeps one record above all others, the state of each step for each student, which is either in progress or secured (Section 21). It also keeps the errors a student makes more than once, because a repeated error is usually a misconception with a name, and a named misconception is something a teacher can step in and resolve.

11.6 The coaching cycle

Each new step is taught through the same six moves. They follow the order that research on explicit instruction and practice supports (Bjork & Bjork, 2020; Rosenshine, 2012), which is also the order a good human tutor uses without thinking about it.

Figure 2

The proposed coaching cycle



Note. The six moves run in order. A missed check sends the step back to Teach in a different form, and a student who is stuck goes to a teacher, not to more of the same problems.

1. **Retrieve.** Five minutes on steps secured earlier, spaced so that older material returns at widening intervals (Dunlosky et al., 2013). Delayed mastery checks happen here too.
2. **Teach.** A short explanation and one worked example, with the reasoning shown line by line. For a step the class has met in Reason, the tutor uses the same model and the same words the teacher used.

3. **Try together.** The student does the next example with support. Help comes as a ladder: first a question that points at what to notice, then a cue, then one worked line. The tutor never opens with the answer, because answers on demand are what harmed learning in the Turkish trial (Bastani et al., 2025).
4. **Practise.** Independent problems at a difficulty where the student gets about four in five right. Rosenshine (2012) found that effective teachers aimed for a success rate of about 80 per cent, and a computational study of simple learning tasks put the best error rate at about 15 per cent (Wilson et al., 2019). Too easy and nothing is learnt; too hard and the student starts guessing. Once each problem type is familiar, the types are mixed.
5. **Explain.** Before a step can be checked, the student says why the method works, in a sentence with "because" in it. In Stage 2 this is a typed sentence from a sentence starter, and in Stage 3 it is a short exchange in which the tutor asks a follow-up question. Explaining is what separates a student who has learnt a trick from one who understands it, and it is the habit Reason depends on.
6. **Check.** Ten fresh problems with no hints. Nine or more correct and the step is marked ready. It is recorded as secured only when the student passes a second, shorter check at least three days later, and a teacher samples those checks in person (Section 21).

Three rules govern what happens when it does not go to plan. If a student is stuck for two minutes, or has climbed the whole hint ladder on one problem, the tutor stops and flags the teacher. If a student makes the same error three times, the tutor names the pattern for the teacher, for example 'treats the equals sign as an instruction to write the answer', and stops serving more of the same problem. If a student misses a check twice, the step goes back to Teach with a different representation: a number line instead of blocks, or an array instead of repeated addition.

The tutor adapts the starting point, the size of the steps, the representation, the reading load, the pace and the amount of review. From Stage 2 it also adapts the context, drawing problems from the student's own week: the compost temperatures from Wednesday's science session, the takings from the market stall, the distance of a training run. It does not adapt the standard. The outcome is the syllabus outcome, the explanation is always required, and nine in ten is nine in ten. Adjustments for students with disability are made by teachers with families (Section 24), not by the software.

A short exchange shows the difference between coaching and answering. A Year 4 student is on the constant difference step, with a problem taken from the Department's unit (NSW Department of Education, 2025).

- **Tutor:** Here is 126 − 58. Do not work it out yet. Which number near 58 would be easier to take away?
- **Student:** 60
- **Tutor:** On the number line, 58 and 126 are this far apart. If 58 slides up to 60, what has to happen to 126 to keep the gap the same?
- **Student:** it goes up 2 as well. 128
- **Tutor:** So 126 − 58 has the same answer as which subtraction?
- **Student:** 128 − 60. 68
- **Tutor:** Tell me in a sentence why that works. Start with "It works because".
- **Student:** It works because both numbers moved the same so the gap didnt change
- **Tutor:** That is the constant difference strategy, and that is a clear reason. Now one with no help: 143 − 97.

The tutor asked three questions, asked for a reason and gave no answers. Had the student typed 124 at the second question, the next prompt would have been a cue about the direction of the slide, not the number.

11.7 The Augment block, minute by minute

An Augment pair runs for 85 to 105 minutes of session time, depending on which pair of sessions it falls in. Table 5 shows the proposed pattern for a Stage 2 class on a day when Augment goes first.

Table 5

A proposed Augment pair in Stage 2 (example Monday)

Time	Students	Tutor	Teachers
9:00	Roll in home groups, then walk to the Stage 2 room		
9:05	Launch: read yesterday's line in the paper learning log and tell a partner today's two goals	Shows each student's map and suggests today's steps	One leads the launch; the other reads the flags left from the last session

Time	Students	Tutor	Teachers
9:08	Retrieve (English): spelling, vocabulary and grammar steps secured earlier	Spaced review; runs any delayed checks that are due	Both circulate
9:13	Cycle 1: Teach and Try together on a new step	Worked example, then the hint ladder	Teacher A pulls a same-wall group of three or four to the table for ten minutes
9:28	Stand, stretch, water	Screen pauses	
9:29	Cycle 2: Practise and Explain	Holds difficulty at about four in five; asks for the because sentence	Teacher B runs three-minute conferences with students whose checks are due
9:40	Check, for students who are ready; others keep practising	Ten fresh problems, no hints	Go to anyone flagged as stuck
9:43	Close: write one line by hand in the learning log (secured, stuck on, or a question for Reason)	Saves the state of every step	Collect the questions for the wall
9:45	Crunch and sip outside		
10:00	Retrieve (mathematics): number facts and last week's steps	Spaced review and delayed checks	Both circulate
10:05	Cycle 1: Teach and Try together	Worked example, then the hint ladder	Teacher B takes a same-wall group; Teacher A conferences
10:20	Stand, stretch	Screen pauses	
10:21	Cycle 2: Practise and Explain	As above, with problems drawn from the class's own data where it can	Circulate; go to flags
10:35	Check, or continue	Ten fresh problems, no hints	Sample two or three checks in person
10:42	Close: learning log, pack up	Sends each teacher a three-line summary of the room	
10:45	Recess		

Note. Of the 105 minutes between 9:00 and 10:45, about 75 are at the screen, 15 are crunch and sip, and the rest is talk, movement and writing by hand. Times are indicative.

11.8 What teachers do during the block

The teacher's screen shows three short lists, not thirty thumbnails of students' screens.

- **Stuck now.** Students flagged by the two-minute rule. A teacher goes to them.

- **Same wall.** Three or more students making the same named error. A teacher pulls them to a table for ten minutes of explicit teaching with materials, which is the small-group lesson on remainders described in Section 17.
- **Ready to check.** Students whose delayed checks are due. A teacher samples them in a three-minute conference: show me one, and tell me why.

Two further jobs never appear on a list. One is watching for the student who is gaming the system. The tutor reports answer patterns that look like guessing or hint-harvesting, such as answers entered faster than the problem can be read, but it does not watch faces, eyes or anything outside the task (Section 25). The other is noticing the student whose map looks fine and whose morning does not.

This is why the adults in Augment have to be teachers. Alpha asks its guides to motivate. These teachers diagnose a misconception from a pattern of errors, teach to it on the spot and judge whether an explanation shows understanding. With about 26 students and two teachers in the Stage 2 room, each teacher can run one small group and three or four conferences in a session. Across the eight English and mathematics Augment sessions of the example week, that is enough for every student to have a teacher conversation about each subject about once a week.

11.9 Motivation without a points economy

Alpha is candid that software is the smaller part of its model and motivation the larger, and it builds motivation from points, a school currency and prizes. The reporting from Brownsville is a warning about where that can lead when targets start to bite (Feathers, 2025). Across more than a hundred experiments, expected tangible rewards for doing a task reduced people’s interest in the task once the rewards stopped (Deci et al., 1999).

The proposed Augment sessions draw on four other sources of motivation, each tied to one of the needs in self-determination theory (Ryan & Deci, 2000).

- **Visible progress.** The student’s map fills in step by step against the syllabus, and 'secured' means something because it was checked twice.
- **Choice.** Students choose the order of the day’s steps and, from Stage 3, set weekly goals with their teacher.
- **A teacher who knows.** The weekly conference, and a home-group teacher who has seen the map.

- **Use.** Within the week the step is needed for something real: the float for the market stall, the graph of the compost data, the letter to the council. This is the loop in Section 15 doing motivational work.

There are no points, no leaderboard and no currency, and nothing a student eats, plays with or is paid depends on a target. Progress is counted in outcomes secured and never in minutes at the screen. A student who finishes the day's steps early reads, or helps a classmate. Whether this is enough to keep an unselected class working hard at a screen for two sessions a day is one of the first things a pilot has to find out, and I would rather find out than assume.

11.10 Guardrails

- **Content.** Every explanation, worked example and check item comes from a bank written or approved by teachers and mapped to the syllabus. Generative AI is used to vary contexts, respond to a student's explanation and choose the next move. It is not used to invent the curriculum as it goes. The reporting on faulty machine-written lessons at Alpha is the reason (Maiberg, 2026).
- **"I think the tutor is wrong."** Every screen has that button. It sends the item to the teacher and it counts as good practice, not cheek. It is digital literacy applied daily (Section 13).
- **Answers.** In practice the tutor does not give final answers, write sentences for students or accept pasted text.
- **Conversation.** No open-ended chat before Stage 2, scripted prompts and sentence starters in Stage 2, and fuller dialogue in Stage 3, always inside the task. The tutor has no name, face or personality and never presents itself as a friend (Section 24).
- **Distress.** Anything that suggests distress or risk goes to the teacher at once (Section 24).
- **Privacy.** No cameras, no attention or emotion tracking, no homework on the tutor, Australian hosting, and conversations open to teachers and families and closed to everyone else (Section 25).
- **A person decides.** The tutor can mark a step ready. Only a sampled check and a teacher's judgement put it in a report.

11.11 How coaching changes from Kindergarten to Year 6

Table 6

Adaptive practice and coaching by Stage (proposed)

	Kindergarten to Year 2	Stage 2 (Years 3–4)	Stage 3 (Years 5–6)
Screen time across the two sessions	About 45 minutes	Typically about 75 minutes, never more than 90	Typically about 75 minutes, never more than 90
Who leads	Teachers in small groups; software for short practice	The tutor for practice; teachers for small groups and conferences	The same, with more goals set by the student
What the software is like	A practice game: tap, drag and listen, with no chat	The coaching cycle with scripted prompts and sentence starters	The coaching cycle with short Socratic exchanges
Typical content	Phoneme–grapheme correspondences, blending, number facts, vocabulary	Additive and multiplicative strategies, place value, spelling, grammar, comprehension	Fractions and decimals, multi-step problems, longer texts, editing
The Explain move	Aloud, to a teacher or a partner	A typed because sentence	A short exchange with a follow-up question
The mastery check	Watched by a teacher	A fresh set and a delayed check, sampled by a teacher	The same, with the student predicting their own result first
Inquiry sessions	None	The tutor helps frame questions	The tutor helps frame and sharpen questions

11.12 What is claimed

Everything in Sections 11.5 to 11.11 is a design hypothesis. The parts have evidence behind them. The assembly does not, for students under twelve, here or at any other school. As far as I know, no tutor available in 2026 does all of this against the NSW syllabuses under the privacy commitments in Section 25. Some of it could be configured from existing tools and some would have to be built. Appendix A applies the method to two real units, which is the first thing I would build and the first thing I would test (Section 30).

12 Reason

Reason is learning with a teacher, taught as a class and with the screens off. In the proposed design every student has two Reason sessions a day, in classes of about twelve, and they are where English, mathematics and science and technology are taught by a person rather than

practised with a tutor. In English that means the part of the subject no machine should be doing for a nine-year-old: reading a real book and talking about it, writing by hand and then redrafting, being read to, arguing about a poem, and standing up and telling the group what happened on the weekend, properly, with a beginning and an end. While the teacher works line by line with six students on their drafts, the other six read. Then they swap. Reason is protected from AI by the timetable rather than by preference, and because it rotates through the day it gets its share of first sessions, which nobody can quietly trade away for something more measurable.

English has the largest share of Reason because writing is where the AI question bites hardest. A chatbot can produce a passable persuasive paragraph in five seconds, which is exactly why a student needs to have produced hundreds of them, slowly, with a teacher reading every line. The evidence on writing instruction in the primary years favours explicit teaching of strategies for planning and revising, frequent feedback and plenty of time spent actually writing (Graham et al., 2012), and handwriting instruction itself improves the quality and fluency of students' writing (Santangelo & Graham, 2016). Reading on paper supports comprehension better than reading the same text on a screen, for students as for adults (Delgado et al., 2018; Furenes et al., 2021). None of this is a rejection of keyboards, which help many young writers (Graham et al., 2012) and which students use elsewhere in the week. It is a decision about what these sessions are for. Talk matters too. Structured classroom dialogue, in which students explain, justify and challenge ideas, is a well-developed approach to building oral language and reasoning (Alexander, 2020).

Augment handles the skill-based parts of the English syllabus, such as spelling, vocabulary, grammar and comprehension practice at the right level. Reason handles meaning: composition with a purpose and an audience, and discussion of texts with people who have read them. The teacher marks the writing with care, sits beside the student for the redraft and holds a standard the student cannot bluff past. AI may help draft feedback or spark ideas, but it does not replace the conversation about what the writing is trying to do and whether it is doing it. Tables 10 and 11 in Section 18.2 set out this division for every focus area in English and mathematics, and Section 18.4 explains how the class unit taught in Reason and each student's own edge in Augment are kept in step.

Mathematics and science work the same way. In mathematics, Reason is where a teacher introduces a new idea to the class, works through problems with them and takes mathematics outside to measure, estimate and collect data, before students practise at their own edge in Augment (Rosenshine, 2012). Science and technology is taught only in Reason,

because it depends on doing things with materials and with other people: running experiments, testing designs and arguing about what the results mean. That is also why the questions raised in Augment come to Reason to be investigated (Section 15). Mathematics and English both have sessions in Augment and in Reason every week, so each student practises at their own edge with the tutor and then works through the ideas with a teacher and the class.

Reason is also where the teacher gets to know the student as a writer and a thinker rather than as a dashboard. Augment gives the teacher data, and Reason gives the teacher the student. Across the example week, English in Augment and in Reason, including the Friday argument, adds up to seven hours, or 30 per cent of the timetable, and the communication work in Create adds more. Mathematics has just over five hours, about 22 per cent (Table 8). The difference from a conventional timetable lies in the mix rather than the quantity. Practice without meaning produces fluent bluffers, and meaning without practice produces students who cannot spell.

In Kindergarten and Year 1, Reason is most of the English programme, joined to the explicit phonics teaching in Augment. Practice can be supported by software, but learning to love a book is a person's job.

13 Digital literacy

Alpha uses AI, but this school also teaches it. In the proposed timetable, every Friday opens with a Reason session of digital literacy, taught by a teacher, with AI as the object of study rather than the instrument. The second Reason session that morning is the Friday argument, where the week's questions about AI are debated without a screen in the room, so together that is about an hour and a half a week, or around sixty hours a year. For comparison, Beijing's new requirement is eight class hours a year (Australian Government Department of Education, 2025). This is also the thing PrimarAI has been doing in NSW classrooms through incursions since 2025, so the strand is not invented from scratch; it turns the incursions into a syllabus. Earlier editions called the strand machine literacy.

The strand is built as a scope and sequence from Kindergarten to Year 6, the way mathematics is. It starts with the question that started all of this for me: can you be mean to a chatbot? Five-year-olds can think about that, and they do. From there it builds through what a model is and is not, why it sounds confident when it is wrong, how to check it against a book, a person, a measurement or a walk outside, what it does with what you type, who owns it and why that matters, where the pictures and voices come from, how to tell a real

thing from a made thing, and how to give it a job, judge the result and send it back. Right through, it teaches when to close the lid and do the thing yourself.

Table 7

Digital literacy: proposed outline scope and sequence

Stage	Big question	Students learn to	Example activities	Main syllabus links
Early Stage 1 (Kindergarten)	Can you be mean to a chatbot?	Tell living things, people and machines apart; notice that voices from devices are made by people; talk about kindness and feelings	Sorting game: alive, machine or pretend; the chatbot question as a story circle	PDHPE, Science and Technology
Stage 1 (Years 1–2)	How does a machine learn?	Understand that machines learn patterns from examples, and that they get things wrong; check answers against a book, a person or a measurement	Train an image classifier on their own drawings with Teachable Machine (Carney et al., 2020), then try to fool it	Science and Technology, Mathematics (data)
Stage 2 (Years 3–4)	Why does it sound sure when it is wrong?	Distinguish confidence from correctness; understand what happens to what you type; tell made images and voices from real ones	Fact-check an AI answer about the local area against the library and a local expert; a privacy audit of a children's app	English, PDHPE, HSIE
Stage 3 (Years 5–6)	Who decides, and who is missing?	Recognise bias in training data; check provenance; give an AI a job, judge the result and send it back; reason about AI, work and fairness	Bias audit using questions about their own suburb; the with-and-without week; provenance checks on images; the Friday argument	Science and Technology, English, HSIE, PDHPE

In the with-and-without week, every Stage 3 group builds the same project with AI and without it, then presents which was better and why. Students learn to check where an image came from, because spotting synthetic images by eye is already unreliable and the durable skill is provenance. The Friday argument is where the week's questions get fought out without a screen in the room: whether the tutor should be allowed to see your writing, whether an AI can be a friend, whether it is cheating if the AI only asked the question.

Honesty is taught through dilemmas elsewhere in the design, and this is the richest source of dilemmas a primary school has had in a long time.

The four domains of the European Commission and OECD framework, engaging with, creating with, managing and shaping AI (European Commission & OECD, 2025), run through

the sequence, and the content maps to the digital technologies content of Science and Technology K–6 and to PDHPE content on safety and relationships. This is the part of the model I am most confident about, because I have watched versions of it work in classrooms. Whether a full seven-year strand works as well as a set of incursions is still a hypothesis. Augment borrows from Alpha's tutoring and Estonia's pedagogy. The five Cs borrow from a long tradition of good schooling that mostly got squeezed out. Digital literacy as a taught subject, with its own time and its own progression, is the part that is most our own, along with the rotating shape of the ARC Day itself.

14 Create and the five Cs

Create is the part of the day where AI mostly stays out of the way. In the proposed day every student has two Create sessions, almost eight hours a week, given to the creative arts, PDHPE and sport, team building, making and enterprise. Music, drama and dance, and visual arts are each taught every week, alongside PE, health and wellbeing, team games and sport, so the creative arts and PDHPE each take a little over 9 per cent of the week. Lunch belongs to the programme too, because on the fortnightly cooking day a lunch that Year 5 made in their Create sessions is not a break from the programme but part of it.

The five Cs, communication, collaboration, creativity, character and connection, are anchored in NSW life. Each is taught explicitly, year on year, with progression and public demonstration. They are where students spend most of their time with other people and real consequences, and where the subjects of the curriculum meet real life. They are also where safeguarding matters most, because they take students into the community (Section 24). The examples that follow show how the proposed programme would run.

Communication
Real public speaking is taught from Kindergarten, when five-year-olds tell a story to the whole school. Year 3 students give a three-minute talk at the local cafe to adults who did not have to come. From Year 4 students produce a fortnightly podcast for families and the community, Year 5 writes to the council and to government inquiries, and Year 6 stages a public speaking event with local media in the audience. Debating, journalism and oral history run through the week. By the time a student leaves, they have stood up in front of a room of adults dozens of times.

Collaboration

Collaboration is taught through real projects with real consequences. Year 5 and 6 teams run a stall at the Bondi Markets on a rotation, with a budget to acquit and customers to please. Once a fortnight a cross-age team cooks lunch for the whole school within a budget. Year 6 students run reading sessions with Kindergarten, the market garden is run by cross-age teams from Years 3 to 6, and Stage 3 students take turns in conflict resolution roles. Working together is practised on real tasks every week, not taught in a workshop once a term.

Creativity

Creativity is not a specialist subject squeezed into ninety minutes a week. It is closer to the medium in which much of school life happens. Every student learns one instrument seriously, with weekly lessons and daily practice. The school stages two productions a year, written, designed and directed partly by the students. Visual art happens in a real studio with a working artist. Making is taught, from woodwork and sewing to ceramics and electronics, and so is fixing: bikes, garden tools, broken furniture and simple appliances. In the old Scandinavian sloyd tradition you give students real saws rather than plastic ones, so they actually learn to cut. Every student by the end of Year 6 has made things, fixed things, performed things and exhibited things, often enough that the habit is properly formed.

Character

Character is taught explicitly rather than left to develop by accident. These dispositions are taught the way you teach a language, small and often, with high expectations and consistent feedback, over years. Persistence is taught through physical challenges that take weeks to prepare for. Earlier versions of this paper set fixed distances for everyone, rising to a half marathon at the end of Year 6. That was wrong on two counts: Sydney's major half marathon does not accept entrants under sixteen (Runaway Sydney Half Marathon, n.d.), and a fixed distance excludes students whose bodies work differently. The revised design keeps the intention, a challenge that takes real training and that the student finishes in public, and sets it with each student and family: a run, walk, ride, roll or swim, scaled so that it stretches that student. Hard work is taught through long projects that do not get easier, including a season-long market garden, a six-month research project and a daily practice log for an instrument. Kindness is taught through service, with regular visits to a local seniors' community and peer support roles for younger students. Honesty is taught through dilemmas, mock trials, ethics circles and the regular requirement to defend a position you do not personally hold.

Connection

Connection runs in five directions, taught with roughly the same seriousness as literacy.

- **To the bush.** Camps that build in complexity from a single overnight in Year 1 to a multi-day wilderness trip in Year 6. Bushcraft, navigation, weather-reading, snake awareness and native plant identification, as a recurring rhythm of the year rather than a once-a-year novelty.
- **To the water.** Swimming and water safety for every student, kayaking and sailing from Stage 2, surf life saving skills, and marine science projects with university researchers as citizen science. Every student leaves meeting at least the national benchmark for their age, which is to swim fifty metres continuously and tread water for two minutes, a benchmark about four in ten Australian twelve-year-olds currently miss (Royal Life Saving Society – Australia, 2020). Students who are ready go further, including supervised open-water swims.
- **To civic life.** Regular visits to council meetings, courts, parliament, workplaces and hospitals. Every Year 5 student sits through a council meeting at least once, and every Year 6 student visits a working courtroom and writes about what they saw.
- **To older Australians.** A sustained relationship with one local seniors' community, with the same students visiting the same residents across several years, oral history projects recorded and given back as gifts, and older people invited into the school to teach gardening, woodwork, mending and storytelling.
- **To students in other places.** An exchange with a primary school in Western Sydney and one in regional NSW, sustained over years, so that students know people whose lives are different from their own.

These are hours we have been quietly taking from students for a hundred years to manage the timetable, and the design gives them back more or less without apology.

Alpha runs something structurally similar in its afternoons, organised around domains such as leadership, public speaking, entrepreneurship and grit. Its published workshops for Grades 4 and 5 include launching a food truck or a short-stay rental, and for Grades 6 to 8 securing US$10,000 in funding for a business and finishing a Spartan Race (Alpha School, n.d.-e). The structural idea is shared; the cultural setting is not. The Australian version foregrounds civic life, the natural environment and a less individualist idea of enterprise. Our students do not chase startup funding. They speak at the local council, run a market garden, handle themselves in the water and acquit a community budget. The idea that a large part of the

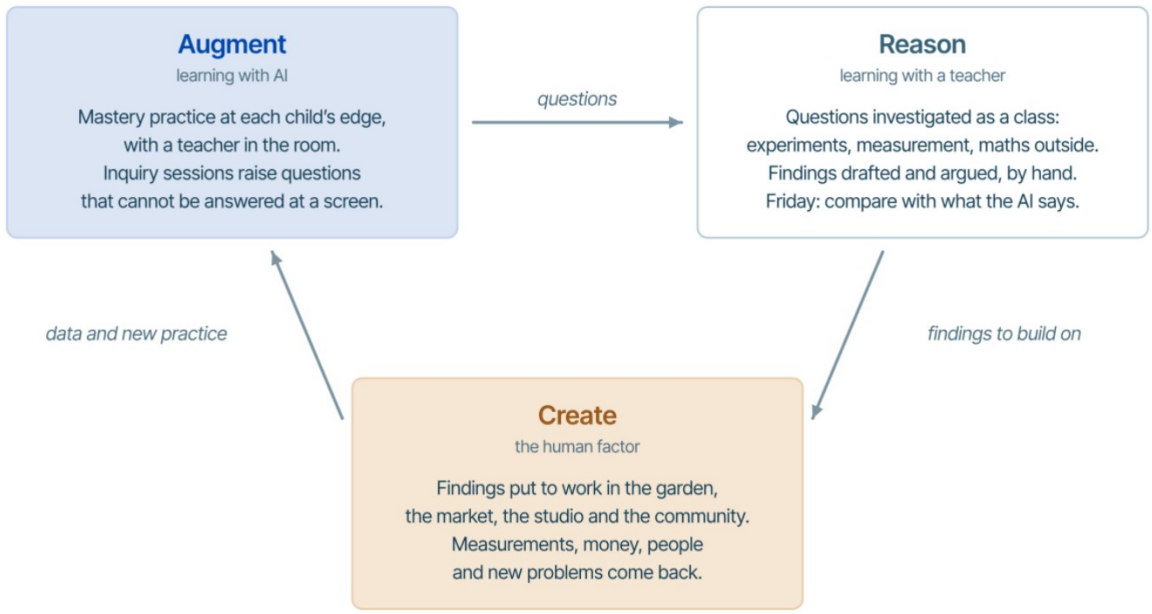
school day could be given over to this was borrowed from Alpha. The shape of the ARC Day, Reason, digital literacy and the substance of the five Cs are decisively local.

15 How the parts make one day

In Alpha's model as publicly described, the academic block and the afternoon workshops run largely in parallel. Its specialty schools make the point. NextGen, Nova and Waypoint change the afternoon and leave the academic block as it is (Section 3.1.2). In the ARC Day the three parts are meant to feed each other, and that loop is what makes the day one piece of work rather than three.

Figure 3

The proposed PrimarAI learning loop



Note. Questions raised in Augment are investigated in Reason and Create; the measurements, writing and problems that come back become practice in Augment and drafting in Reason.

Take, as an example, the compost question. On Wednesday, in an Augment mathematics session, a Stage 2 class's questions include why bed three runs hotter than bed one. That afternoon the class takes the question into its science and technology session in Reason, measures the temperature of every bed with probes and records it in a spreadsheet. On Thursday morning the tutor sets the same students problems in averages and graphs using their own data, and in Thursday's long-writing session in Reason they draft a science report that the teacher reads line by line. On Friday the digital literacy session asks an AI the same question and compares its confident answer with the measurements, and in Monday's

making and enterprise session the garden team builds a better compost heap. Each part of the day has done the job it is best at.

16 The school year

The school year keeps NSW's four terms, because families have jobs and NESA has forms. Underneath them the programme follows the natural year in Sydney, marked by what is flowering, what is moving and what the weather is about to do. The market garden plants and harvests by it. Camps go out in late winter and early spring, when the bush is dry and the snakes are slow, and not in early summer when the storms come. The water programme starts when the water warms and stops when the bluebottles arrive. Each term opens with a whole-school gathering and closes with students presenting what they grew, swam, built or made, so that four times a year the school stops and looks around.

17 An example Wednesday at the school

This is an ordinary Wednesday in the proposed school, the kind a visiting parent or a curious principal might see if they stayed until home time. From the outside the school looks like any of the older small schools of the Eastern Suburbs. Inside, the difference is in how the day is shaped rather than in the fittings, although there is a working market garden out the back and the hall doubles as a theatre.

Most students walk themselves in by Year 2. At 9:00 the multi-age home groups gather for the roll, run by a different Stage 3 student each week, with a thought for the day and a moment of quiet.

Wednesday is a Create-first day, so it starts with the arts. The first session is music, with each class rehearsing with the resident musician while everyone is fresh. After crunch and sip the second is health and wellbeing, PDHPE in small groups, from naming feelings in Kindergarten to staying safe online in Stage 3.

Recess is half an hour on the grass, with no phones, ever. It is the only entirely unstructured part of the day.

At 11:15 the school moves into Augment, English first and then mathematics, in three rooms, one for each Stage group. In the Stage 3 room Mia, in Year 6, is working through MA3-MR-01, strategies for multiplication and division. Her tutor noticed two weeks ago that she had been quietly bluffing her way through long division since Year 4, and she is now backfilling, fast, with attention she would never get in a class of twenty-nine. She is in her second focus

cycle, has just typed a sentence explaining why her method works, and will sit a ten-problem check tomorrow. One of the room's two teachers is running a small-group lesson for three students who all hit the same misconception about remainders, while the other pulls Mia aside to talk about the writing she has been drafting. In the Kindergarten to Year 2 room a Year 1 boy has just finished ten minutes with a teacher and four other students on vowel digraphs for EN1-PHOKW-01, and is now doing a few minutes of phoneme-to-grapheme practice on a tablet. By 12:40 the Stage 2 room has a wall of questions about compost temperature ready for the garden. Nobody is teaching from the front of a room of thirty, which is what visiting principals notice first, and the rooms are quiet in the way a small library is quiet.

Lunch is sausage rolls and a green salad with vegetables from the garden, cooked by a Year 5 team who planned the menu, kept to the budget and did the preparation with the kitchen teacher in Monday's making and enterprise session. The whole school eats outside.

At 1:15 the screens go off across the whole school for Reason. In one Stage 2 group a teacher is going line by line through the interview questions six students will use at the seniors' community in Surry Hills tomorrow, where they will ask residents what childhood was like in the 1950s, while the other six read. A Stage 3 group is redrafting, by hand, persuasive letters to the council about a missing pedestrian crossing. Kindergarten is in a story circle on the rug, retelling a story in their own words before being read to. At 2:00 the Stage 2 class takes the compost question into the garden with probes and a spreadsheet for science and technology.

Home time is 3:00, with extended care until 5:30 that is closer to a bushcraft and music studio than a holding pen. There is no routine homework, because the project work most schools send home is done here, in Create, with other people and out in the world. On other days the cross-age teams from Years 3 to 6 use their Create sessions to build beds, cost seedlings and plan the parent fundraiser, and Year 2 works on woodwork projects they will take home.

A Friday looks different. Reason goes first, with digital literacy, so the Stage 3 class might spend the session getting an AI to write the letter to the council, laying it beside the ones written by hand on Wednesday and arguing about which one the mayor would actually read. A fortnightly Tuesday looks different again, because the whole school goes off-site, to the Royal National Park in autumn, the harbour in spring, a courthouse on a sitting day or the kitchen of a working restaurant midweek. Once a term there is an overnight camp, and once a year Stage 3 goes on an extended wilderness trip in the Blue Mountains while Stage 2

spends a week with its exchange school. The six hours of the school day are real education, not preparation for an education that happens somewhere else later.

Part V: The school in practice

Part IV described the day. This part works through what it would take to run it as a registered NSW school: curriculum, grouping, teachers, assessment, graduation, partnerships, inclusion, safeguarding and governance. Where the original design made claims that do not survive scrutiny, I say so and show the revision.

18 Curriculum and the NSW syllabuses

The school is designed to be registrable from day one as a non-government school with NESA. Registration requires the school to teach the NESA syllabuses for the six primary key learning areas, English, Mathematics, Science and Technology, HSIE, Creative Arts and PDHPE, and to keep the documentation that shows it does (NESA, 2026). New K–6 syllabuses for Creative Arts, HSIE, PDHPE and Science and Technology become mandatory in 2027, alongside the English and Mathematics K–10 syllabuses already in place (NESA, n.d.-a). A school opening in 2027 or later would be built on the new syllabuses from the start.

NESA's K–6 time allocation advice suggests half of teaching time for English and Mathematics together, with at least 20 per cent for Mathematics, 8 per cent each for Creative Arts, HSIE, PDHPE and Science and Technology, and 18 per cent for languages and other activities (NESA, n.d.-b). The advice is a guide rather than a rule, but it is a fair test of whether a new timetable is honest about what it teaches. Table 8 shows the proposed ARC week against it.

Table 8

Weekly teaching time by learning area in the proposed ARC week

Learning area	Where it is taught	Hours per week	Share of week	NESA suggestion
English	Augment (4 sessions) and Reason (5 sessions, including the Friday argument)	7.0	30.0%	50% for English and Mathematics together
Mathematics	Augment (4 sessions) and Reason (3 sessions, including mathematics outside)	5.1	21.8%	At least 20%
Science and Technology	Reason (2 sessions, including digital literacy)	1.8	7.5%	8%
HSIE	Augment inquiry sessions (2)	1.7	7.1%	8%

Learning area	Where it is taught	Hours per week	Share of week	NESA suggestion
Creative Arts	Create: music, drama and dance, visual arts	2.2	9.3%	8%
PDHPE	Create: PE, health and wellbeing, team games	2.2	9.3%	8%
Other activities	Create: assembly and team building, making and enterprise, community and showcase, sport	3.5	15.0%	18%
Total		23.3	100%	

Note. The teaching week is 1,400 minutes: six sessions a day between 9:00 and 3:00, excluding crunch and sip, recess and lunch. Figures come from the timetable in Table 4, rounded, and would be audited against programmed lessons. Mathematics, science and HSIE also appear in Create projects and off-site days, which are not counted here.

The September edition of this paper exposed a problem with the two, one, three day. Mathematics reached 20 per cent only if about two hours a week of it were deliberately programmed into afternoon projects, through budgeting, measurement, data and costing, and without that it fell to about 13 per cent. The ARC Day solves this in the timetable rather than in the planning, because mathematics now has its own sessions in both Augment and Reason. English and Mathematics together take just over half the week, close to NESA's advice, and every other learning area sits within about a percentage point of 8 per cent: Science and Technology and HSIE slightly under, Creative Arts and PDHPE slightly over.

Physical activity needs separate attention. NSW government schools must plan 150 minutes of moderate to vigorous physical activity a week, including 60 minutes of sport or physically active PDHPE in Years 3 to 6 (NSW Department of Education, 2024). As a non-government school PrimarAI would not be bound by that policy, but it is a sensible benchmark. The example ARC week timetables 145 minutes of PE, team games and sport, and the daily movement breaks between Augment sessions and the physical work of Create add to it.

The general pattern for every learning area is the one worked out for English in Section 12: AI for the practice, people for the meaning. The PrimarAI Blueprint, to be published in stages, will work through Mathematics, Science and Technology, HSIE, PDHPE and Creative Arts in the same way, outcome by outcome. Sections 18.1 to 18.4 set out the method, and Appendix A is the first instalment.

18.1 From outcome to evidence

Time allocation shows that a subject is on the timetable. It does not show how a student comes to meet an outcome. Fulfilling a syllabus means four things for every outcome: it is taught, it is practised, it is used, and there is evidence that the student can do it. In a conventional class one teacher does all four inside one sequence of lessons. The ARC Day gives each of the four a home.

Table 9

How one syllabus outcome travels through the proposed ARC Day

What happens	Where	Who leads	What it leaves behind
Taught	Reason, and teacher small groups in Augment	A teacher	A dated entry in the teaching program
Practised	Augment	The tutor, supervised by teachers	The mastery record: each step in progress or secured
Used	Reason and Create	Teachers, with real audiences	Work samples: writing, measurements, budgets, performances
Evidenced	All three	Teachers	Sampled checks, moderated work samples, the portfolio and the A to E report

18.2 Which part of the day carries which part of the syllabus

Tables 10 and 11 divide the focus areas of the English and Mathematics syllabuses across the three parts of the day. The pattern is the one in Section 12. Number and the skill-based parts of English suit coached practice. Anything that depends on materials, talk, a whole text or a real audience belongs to people.

Table 10

English K–6 focus areas across the proposed ARC Day

Focus area	Augment	Reason	Create
Oral language and communication	Small-group talk with a teacher in Kindergarten to Year 2	Class discussion, turn and talk, presentations, the Friday argument	Public speaking, the podcast, interviews with older Australians
Vocabulary	Spaced practice of taught words; words in context	Words met in the mentor text and taught explicitly	Subject vocabulary used in projects

Focus area	Augment	Reason	Create
Phonological awareness, print conventions and phonic knowledge (to Stage 1)	Teacher small groups, then short adaptive practice	Shared and decodable reading	Not taught here
Reading fluency	Short re-reading with a teacher in Kindergarten to Year 2	Reading aloud, paired reading, being read to	Reading to Kindergarten buddies; performance
Reading comprehension	Practice at the student's level: literal, inferential, vocabulary in context	Close reading and discussion of whole texts on paper	Reading for a purpose: council papers, recipes, instructions
Creating written texts	Sentence-level skills: grammar, punctuation, sentence combining	Planning, drafting by hand, feedback and redrafting	Writing for real readers: letters, scripts, reports
Spelling	Adaptive, spaced practice	Taught patterns and morphology; proofreading own work	Used, not taught
Handwriting, and digital transcription from Stage 2	Handwriting in teacher small groups to Year 2; keyboarding practice from Stage 2	All drafting by hand	Used, not taught
Understanding and responding to literature	Not taught here	Mentor texts read and discussed as a class	Performing, illustrating and publishing responses

Table 11

Mathematics focus areas across the proposed ARC Day (Stage 2 names)

Focus area	Augment	Reason	Create
Representing numbers using place value	Coached practice: partitioning, ordering, renaming	Introduced with materials and place value charts	Budgets and takings
Additive relations	Strategy practice, fluency, missing values	Strategy talk, games, word problems worked as a class	Float, change and costing for the stall and the kitchen
Multiplicative relations	Fact fluency with spacing; fact families	Arrays with materials; number properties	Scaling recipes; seedlings per bed
Partitioned fractions	Coached practice on models and number lines	Folding, cutting and comparing with materials	Recipes; sharing a harvest
Geometric measure	Short consolidation: units, conversions, grid references	Measuring and mapping outside	Garden beds, woodwork, navigation on camps

Focus area	Augment	Reason	Create
Two- and three-dimensional spatial structure	Short consolidation	Hands-on work with shapes, nets and models	Woodwork, sewing, ceramics
Non-spatial measure	Short consolidation: reading scales and clocks	Measuring mass, capacity and time with real instruments	Cooking; timetables for off-site days
Data	Reading and interpreting graphs, using the class's own data	Collecting and representing data from real questions	Compost temperatures; stall sales
Chance	Short consolidation	Experiments with spinners and dice	Games
Working mathematically	The Explain move in every cycle	Reasoning and communicating in every lesson	Problem solving under real constraints

18.3 Programming, and what NESA would see

The school keeps the documentation NESA expects of any registered school (NESA, 2026), which in practice means a scope and sequence for each learning area and Stage, a teaching program for each unit, timetables, assessment records and samples of student work. The tutor adds one document a conventional school does not have, an outcome-by-outcome mastery record for every student, exportable at any time.

What changes is how a unit is written. A conventional unit is a list of lessons. An ARC unit keeps the same outcomes, content points and texts and adds one column: for each content point, which part of the day teaches it, which practises it and which uses it. Appendix A does this for two units published by the NSW Department of Education, and nothing in either unit is dropped. The English unit's ten 40-minute lessons map one to one onto the ten Reason English sessions of a fortnight, and its skill-based content is practised in Augment alongside phonics, spelling and handwriting. The mathematics unit's eight lessons are split down the middle. The teaching and discussion in each lesson go to Reason, three sessions a week, and the practice goes to Augment, four sessions a week.

18.4 The class unit and the personal edge

There is a real tension between mastery pacing and a shared curriculum, and the design should say how it resolves it. If every student moves at their own pace through everything, there is no class left to read a book with. If the class moves together through everything, there is no mastery. The proposed answer is two lanes.

- **The unit lane.** Reason runs class units on the Stage scope and sequence at a shared pace. Everyone in Stage 1 reads the same mentor text in the same fortnight, and everyone in Stage 2 meets the constant difference strategy in the same week. In Augment, about a third of each session practises what the class met in Reason that week, at a difficulty set for each student.
- **The edge lane.** The other two-thirds of Augment is the student's own edge on the syllabus map: backfilling a gap from an earlier Stage, securing this Stage's steps, or extending beyond them.

The proportions are a starting point, and a teacher can change them for any student. A student well behind in number may spend most of their mathematics Augment time in the edge lane for a term and still take full part in the class's Reason lessons, with scaffolding. A student who has secured the Stage's outcomes extends within the Stage first, through harder problems and applications, and moves into the next Stage's content with the teacher's sign-off and the family's knowledge. That follows the Stage structure NESA already uses and the argument of the NSW Curriculum Review for progression when ready (Masters, 2020).

The cost is that a student's Reason lesson and Augment edge will sometimes be far apart. That is already true in every classroom in NSW (Goss et al., 2016). The difference here is that the gap is visible, named and worked on daily. Whether one-third and two-thirds is the right split is a design hypothesis for the pilot.

19 Progression, grouping and home groups

Academic progression in Augment is mastery-based, within each student's NSW Stage group. The cohort matters socially, so cohorts stay intact for many purposes, but the cohort is not where academic progression happens. NESA's Stage structure already groups two years together, which makes this easier than it sounds. One earlier claim needs correcting. The mastery approach to mathematics associated with Singapore and Shanghai is a whole-class method that keeps students together while securing understanding, which is different from individual pacing.

Home groups are different. They are multi-age groups of about twelve students from Kindergarten to Year 6, with one teacher, and they exist for relational rather than academic reasons. The research on multi-age classes finds no meaningful difference in achievement compared with single-age classes and, if anything, small benefits in attitudes and social outcomes (Veenman, 1995). That is a useful result. Multi-age grouping does no academic

harm and is not academic magic, so it should be justified by what it does for belonging, peer teaching and leadership, which is the point here.

The original design said the home-group teacher would stay with the same students across their whole primary journey. There is good evidence that being taught again by a teacher who already knows you helps (Hill & Jones, 2018), but seven years is a long time to be matched with one adult. The revised design aims for a teacher to stay with a group for several years, ideally a whole Stage or longer, reviews every match each year with families and gives a student a straightforward way to move groups if the relationship is not working.

20 Teachers

Far from disappearing, the teacher's role in this model is elevated. It becomes the work a chatbot has never done well and most likely will not: reading a student's face, holding a high standard when they want to give up, building character and making a class feel like a family. Alpha's teachers have said, almost without exception, that they are finally teaching the way they had hoped to and having an effect they never had in a conventional school. The Castlereagh Statement's language is close to the right shape. It calls for teaching to be redefined "as a profession centred on what matters because of AI: nurturing identity, cultivating purpose, building the relational and metacognitive capabilities that enable learners to thrive beyond technology" (Castlereagh Summit, 2026, Appendix B). That is more or less what I came into teaching to do.

The people who staff this school are always called teachers. Alpha employs guides rather than licensed teachers (Sands, 2026), and that is one of the clearest differences between the models. Alpha takes its guides seriously. It reports more than 80,000 applicants for a few hundred places, says about half of its guides hold a master's degree or doctorate, starts them on US$100,000 and gives each a group of students to answer for (Alpha School, n.d.-a, 2026). But it selects them to motivate and to run workshops, not to teach, and the head of its Brownsville campus has said that guides do no teaching (Feathers, 2025). Section 11.8 sets out why Augment needs the opposite. Every adult leading learning at PrimarAI is NESA-accredited, because supervision in Augment, line-by-line feedback in Reason and judgement about risk and growth in Create are all professional work.

The workload case is urgent. Australian lower-secondary teachers reported a 46.4-hour working week in the OECD's 2024 Teaching and Learning International Survey, against an OECD average of 41, and 69 per cent named administrative work as a major source of stress (OECD, 2025b, as reported by Cobbold, 2025). Grattan Institute has argued for years that

teachers need more time for the work that matters most, such as planning, feedback and collaboration (Hunter et al., 2022). More than anything else in the design, this structure is meant as a rescue of teaching as a serious adult profession. The teachers I would want at PrimarAI are exactly the people who are leaving because the job has become photocopier-jockey for a curriculum that AI has already disrupted underneath them. In this design Augment is also where teacher time is released: while some teachers supervise and teach small groups, others are reading writing, planning projects and writing narrative reports inside the school day.

The rollout is teacher-first, as Estonia's was. No student meets the system before their teacher has spent a term inside it, learning how the tutor behaves, where it fails and how to override it. Alpha trains a new guide in a two-week cohort before giving them a group (Alpha School, 2026). A term is a fair measure of how much more this design asks of its teachers. A teachers' programme of school-based learning communities, shared materials and work with subject teacher associations runs alongside.

20.1 The staffing equation

The September 2026 version of this paper claimed that "the economics of fewer staff at higher pay actually does work." On my own numbers it does not, at least not in the form stated. Home groups of twelve, Reason classes of twelve and Create sessions of small, supervised projects need more teachers per student than a conventional primary school, not fewer. Table 12 shows an illustrative model.

Table 12

Illustrative staffing for a school of 168 students (24 per year, Kindergarten to Year 6)

Role	FTE	What the role does
Home-group teachers	14	One per home group; also teach Reason, lead Create project teams and write narrative reports
Specialist teachers	3	Music, visual arts and making, PDHPE and outdoor learning
Learning support teacher	1	Works across the Augment rooms and with families
Principal and deputy principal	2	Both teach part-time; the deputy leads assessment, moderation and mastery data
Kitchen and garden coordinator	1	Supervises cooking and the market garden
Learning technology and data protection lead	0.5	Vendor management, privacy, audit and incident response
Administration and extended care	2.5	Office, enrolments, extended care

Note. In each Augment session, about six teachers work in the Kindergarten to Year 2 room and two in each of the Stage 2 and Stage 3 rooms, with the learning support teacher moving between them, which leaves about four home-group teachers released on rotation.

That is about twenty teaching staff for 168 students, nearly twice the teaching staff of a conventional primary school of the same size, and paying them more than the public scale compounds the cost. The model could be made cheaper by running larger Augment rooms, increasing school size or reducing Create, but each of those changes something educationally important. PrimarAI, as designed, is a high-staffing model whose costs must be met from a mix of government recurrent funding, fees, philanthropy and scholarships, and a full financial model is one of the first pieces of work required before launch (Section 32). If the model cannot be made affordable without abandoning small human groups, that is a finding the design has to face.

21 Assessment in six layers

The Castlereagh Statement calls for assessment that protects "both the integrity and relevance of education" and for AI detection to be phased out (Castlereagh Summit, 2026, Principle 3). I agree about detection, which is a losing game when AI text is already hard to tell from human writing. Judgement is the job. But "integrity" is doing a lot of work in that sentence. If assessment is still mostly about a product, and if the product can now be made by AI, then the process matters more, and the only reliable way to see process is for a person to watch it happen over time. Assessment also serves sorting and status in Australian schooling, and a primary school has some freedom from that pressure that it should use. Recent conversations about AI and assessment, including Dylan Wiliam's on the Ed-Technical podcast, have turned towards oral and conversational assessment as ways of seeing learning that a machine cannot fake (Hills & Henkel, 2025).

The harder question is why process should matter so much when AI can now perform it. None of the current language quite fits. What is needed is a layered system that tells families, teachers and the student what the student is becoming, what they can do and where they are heading next. The PrimarAI School assesses in six layers, each designed to be honest, useful and hard to fake.

Table 13

The six layers of assessment

Layer	What it tells us	How it works	Main safeguard
1. Mastery, not marks	What the student has learnt	Each Stage outcome recorded as secured or in progress, with delayed checks	Teacher sampling of mastery checks; comparison with external tests
2. Real performance	What the student can do in the world	Talks, performances, projects and budgets in front of real audiences	Shared criteria and moderation across teachers
3. Portfolio	How far the student has come	A body of work curated by the student, reviewed in conferences	Includes drafts and process, not only final products
4. Self-assessment	How well the student knows themselves as a learner	From Year 2, students assess their own work against criteria	Used for learning, never for grades
5. Narrative portrait	Who the student is becoming	A substantial written narrative from the teacher twice a year	Second reader for every report; evidence cited
6. Graduation sign-off	That the student is ready for what comes next	Three adults sign off the graduation commitments	Clear standards and inclusive pathways (Section 22)

Mastery, not marks. In Augment there are no percentages. A skill is either secured or in progress, and the student keeps working until it is theirs. Every Stage outcome is tracked individually, so families always know which capabilities are secured and which are next. Formative assessment, using evidence of learning to adjust teaching, is one of the best-supported practices in education (Black & Wiliam, 1998), and mastery learning builds it into the structure of the day. Earlier versions of this paper called mastery learning "the most evidence-supported approach in education research." That overstates it. The evidence is positive and substantial, and it is also uneven (Kulik et al., 1990; Slavin, 1987). In practice a step is marked ready at nine in ten on a fresh ten-problem check and secured after a second check at least three days later, with teachers sampling those checks in conference (Section 11.6). Alpha's gate is about 90 per cent on the day (Alpha School, n.d.-a). The delayed check and the teacher's sample are what this design adds.

Real performance and the portfolio. In Create students demonstrate what they have learnt to real audiences, from a three-minute talk at the cafe in Year 3 to a community fundraiser run start to finish in Year 6, and the assessment is the doing itself. Across the primary years each student also curates a portfolio that captures growth, range and the drafts behind the best work.

Self-assessment. From Year 2, students learn to assess their own work. Can they see their own errors? Can they explain why a piece of writing is strong or weak? Can they tell the difference between a confident answer and a correct one? Self-assessment improves learning when it is taught and used formatively, and students' accuracy improves with practice (Andrade, 2019). That capacity is close to the centre of what it means to be a thoughtful adult in the AI era.

The narrative portrait and the sign-off. Twice a year every student receives a substantial written narrative from their teacher, more portrait than report card, describing who they are becoming as a learner and reporting on dispositions such as curiosity, kindness, honesty, perseverance and courage. Reporting on dispositions carries a real risk of bias, so every narrative is read by a second teacher and grounded in specific evidence. At the end of Year 6 three adults sign each student off: their home-group teacher, a respected member of the community and an external practitioner in an area of the student's growing expertise. It is one of the oldest forms of assessment there is, an experienced person judging a young person ready to move on.

Two compliance points need stating. Schools receiving Commonwealth funding must give families plain-language written reports at least twice a year that grade achievement on a five-point scale, which in NSW is conventionally A to E (Australian Education Regulations 2023). The mastery record therefore has to be translated into that scale, which is straightforward if the translation is agreed in advance and explained to families. And PrimarAI students sit NAPLAN in Years 3 and 5. The earlier claim that they would "perform well at it" was a prediction. NAPLAN results are one of the comparison measures in the evaluation (Section 30), and the school will publish them.

22 Graduation commitments

Every Year 6 leaver should be able to show that they have done the fifteen things in Table 14. The earlier version described these as requirements without which "a child does not graduate." That needs two corrections. First, no primary school credential can or should delay a student's move to Year 7, so these are a school credential, not a gate. Second, several of the original requirements, particularly the physical ones, would have excluded students with disability or simply set targets outside a student's control. The revised commitments keep every original intention and add an inclusive pathway to each. They are commitments the school makes to every student as much as demands it makes of them. Where a commitment has not yet been met, the record says what was achieved and what comes next.

Table 14

The fifteen graduation commitments

	Commitment	Inclusive pathway or revision
1	Read fluently, write a researched piece of about 2,000 words and edit their own work	Length, format and assistive technology adjusted to the student; dictation and multimodal equivalents accepted
2	Show secured mastery of the Stage 3 Mathematics outcomes	Or of personalised learning goals where adjustments have been made under the Disability Standards
3	Deliver a ten-minute talk to an unknown adult audience	Notes, augmentative communication, a smaller audience or a recorded talk where needed
4	Lead a group discussion on a topic of their choosing	Supported roles and preparation scaled to the student
5	Meet the national water safety benchmark and complete supervised open-water and small-boat sessions	Revised from a one-kilometre open-water swim, which becomes an extension for students who are ready; adapted programmes for students with disability
6	Complete a personal endurance challenge trained for over at least a term	Revised from a half marathon; set with the student and family, in any form of movement
7	Cook a meal for about thirty people from a planned menu and within a budget	As part of a team, with roles matched to the student
8	Plan, run and acquit a community project with a budget of at least $500	As part of a team
9	Play an instrument or sing, with sustained practice over at least two years, and perform publicly	Revised from "intermediate level", which is hard to define fairly; adapted instruments where needed
10	Make at least one substantial physical object	Tools and supports adapted to the student
11	Direct an AI system to do useful work, verify its output and explain when not to trust it	Assessed in the student's strongest mode of communication
12	Identify ten native plants in the local area by name	Any mode of identification
13	Take part in a sustained intergenerational programme over several years	Revised from "hold a sustained relationship," which cannot be required of either party; with consent, safeguarding and support if a resident dies
14	Write to an elected representative about a real issue and follow it through	Revised from "elicits a real reply," which is outside the student's control
15	Articulate who they are becoming as a learner	In writing, speech or another medium

These are fifteen concrete, demonstrable things, judged by people outside the classroom. A parent reading the list knows what they are signing up for, a principal knows what they

would be running and a student knows what is expected and, more importantly, why most of it is worth doing.

23 Families, community and partnerships

A school built around real audiences, real customers and real older Australians cannot run without its community, and partnerships like these only work when they are planned, funded and supported properly. The Castlereagh Statement makes a related point when it asks families and communities to share responsibility for children's healthy relationships with technology (Castlereagh Summit, 2026, Appendix B).

Families see everything the system knows about their child, in plain language, and are invited into the school's work as visitors, teachers of skills and audiences. A family programme on AI at home covers the risks of AI companions and the context of Australia's social media minimum age, which since December 2025 has required platforms to prevent under-sixteens from holding accounts (*Online Safety Amendment (Social Media Minimum Age) Act 2024*). Partners include the seniors' community, the council, local markets, universities, the exchange schools and youth organisations such as debating societies, surf life saving clubs and theatre companies.

Partnerships with students need safeguards. Every adult who works with students holds a Working With Children Check, the school implements the NSW Child Safe Standards (*Children's Guardian Act 2019* (NSW)), and arrangements such as exchange homestays need careful screening, consent and supervision. These safeguards are the conditions under which a community can be trusted with its children.

24 Wellbeing, inclusion and safeguarding

The question of students the model does not fit is the most important in the whole paper, and the one Alpha has handled worst. The answer has to be structural rather than aspirational. Under the *Disability Standards for Education 2005*, the school must consult families and make reasonable adjustments so that students with disability can take part on the same basis as their peers. Universal design for learning provides a framework for building flexibility into goals, materials and assessment from the start (CAST, 2024). Adaptive practice can help some learners, including students who need to work well below or above their age level and twice-exceptional students, but it is not a substitute for specialist support. The learning support teacher, smaller human groups, family partnership and the inclusive pathways in Table 14 are what make the model work for students who are not its easiest fit.

A scholarship pathway from day one helps prevent the school from selecting only engaged, middle-class families, which is precisely the selection effect that makes Alpha's results so hard to interpret.

Screen time is taken seriously. In the proposed day, students in Kindergarten to Year 2 spend about forty-five minutes on screens, across their two Augment sessions. Stage 2 and 3 students spend up to about ninety minutes a day on tutored practice, in two sessions with a break or a change of subject between them, and little screen time in Reason or Create. In the minute-by-minute plan in Section 11.7 the figure is closer to 75 minutes, because each session opens and closes away from the screen. That is educational screen time, not the recreational screen time that national guidelines recommend limiting to two hours a day for children aged five and over (Australian Government Department of Health, 2021), but the school still treats daily totals as something to coordinate with families. The screens are tools, not feeds: a closed system, no notifications, no social media.

The tutor is designed never to present itself as a person or a friend, and it is not a counsellor. It has no camera on the student and does not track attention. If a student says something to it that suggests distress or risk, it tells the student to talk to their teacher and alerts the teacher straight away. Conversations are reviewed by teachers, not by vendors, and younger students do not have open-ended chat at all.

25 Governance, privacy and data

Privacy has to be central, not bolted on later. Consumer AI dropped into primary classrooms and collecting data on six-year-olds is not acceptable. Purpose-built Australian infrastructure has a real cost that governments and all three school sectors need to meet, and NSWEduChat shows it can be done.

Four commitments are written into the school's charter rather than just its privacy policy.

1. **Students' data stays in Australia**, hosted in Australian data centres under Australian law.
2. **No model is trained on students' work or conversations**, and this is a contractual condition for every vendor.
3. **Families can see, in plain language, everything the system holds on their child**, and can ask for anything not required to be kept by law to be deleted.
4. **An independent privacy audit is published every year.**

The revised design adds five more: collecting only the data needed for learning; no emotion recognition, biometric analysis or behavioural monitoring of students, however ambient AI develops; a published incident response plan consistent with the national data breach scheme; human review of any automated decision about a student; and procurement standards that favour tools with evidence of educational effect, echoing Castlereagh's call to reduce reliance on vendors to define good learning (Castlereagh Summit, 2026, Appendix B). Following Estonia's example, students' tutoring conversations are treated as confidential learning records, open to their teachers and families and closed to everyone else (Thian, 2026). This edition adds two more, in direct response to reporting on Alpha: the tutor is not used at home or for homework, and no camera or attention-tracking software watches a student at school (Feathers, 2025).

The legal frame is the *Privacy Act 1988* (Cth) and the Australian Privacy Principles, strengthened by the *Privacy and Other Legislation Amendment Act 2024* (Cth), which requires a Children's Online Privacy Code to be registered by December 2026 (OAIC, 2026). Whether that code will cover school learning platforms is not yet clear. The school will meet its standard regardless. Health information is also governed by the *Health Records and Information Privacy Act 2002* (NSW).

Trust has to be sequenced, not assumed. That is part of why the rollout is teacher-first, why the charter commitments are public and why the school's board includes an education advisory committee with academic, parent and student voices and an AI and data committee with the power to switch any tool off.

26 Place, community and an open model

This is an Australian school, and it should feel like one. Its calendar follows the Sydney year, its Create sessions and off-site days take students into the bush, onto the water and into council chambers, and its idea of enterprise is a market stall rather than a startup pitch. It is public-school adjacent in its values even though it would be registered as a non-government school.

It is also open. The whole design is published under a Creative Commons Attribution licence, so any school can adopt any part of it, whether that is the ARC Day, the digital literacy strand or the six assessment layers, without adopting the rest. That is how I want PrimarAI to work: free to read, free to argue with and free to copy. If the risk of widened inequality comes partly from only students in wealthy families getting AI tutors, AI literacy and days like this,

then publishing openly is part of the response. Publishing can build the theory, but action and money have to do the rest.

Part VI: Evidence, uncertainty and implementation

A manifesto should be ambitious enough to make a proposal and confident enough to let that proposal be tested. This part sorts the claims in the paper by the strength of their support, reviews the evidence for the main components of the design, answers the strongest objections, sets out how a pilot would be evaluated and says in advance what would count against the model.

27 Five kinds of claim

Table 15

The status of the main claims in this paper

Category	What it means	Examples from this paper
Established research	Substantial, reasonably consistent evidence	Well-run tutoring improves learning; formative assessment works; explicit phonics teaching is central to early reading; teacher–student relationships matter; children comprehend better on paper
Emerging evidence	Early, mostly from older students, moving fast	Supervised or guarded AI tutoring can produce strong short-term gains; unrestricted AI help can reduce learning; metacognitive laziness
Reasoned extrapolation	Established findings applied to a new setting	Tutoring and AI evidence applied to primary students; time-of-day and fatigue research applied to a rotating primary timetable; project-based and outdoor learning evidence applied to the five Cs; AI literacy frameworks applied to a Kindergarten to Year 6 strand
Design hypothesis	Original PrimarAI proposals to be tested	The ARC Day produces at least equivalent academic growth while freeing time for the human programme; rotating the first session spreads its benefit across the three parts; Reason improves writing; digital literacy builds durable judgement about AI; the six layers are valid and workable
Forecast	Predictions about technology, work and society	Timelines for human-level AI; ambient AI; changes in work; the end of the agent phase

The design depends on the first three rows and is tested through the fourth. It does not depend on the fifth.

28 What the evidence says

28.1 Tutoring, human and artificial

Benjamin Bloom (1984) famously reported that students tutored one-to-one performed two standard deviations better than students taught conventionally, and challenged educators to find group methods as effective. The two-sigma figure has been quoted ever since, often by people selling software. It came from small, short studies using tests closely aligned to what was taught, and modern estimates are much lower (von Hippel, 2024). A meta-analysis of 89 randomised trials found an average effect of tutoring of about 0.29 standard deviations, largest for programmes using teachers or paraprofessionals, in the early grades, during the school day and at least three times a week (Nickow et al., 2024). That is a large effect by the standards of real-world education interventions (Kraft, 2020), and it is the realistic benchmark for what tutoring of any kind can be expected to do.

Computer-based tutoring has a long record of its own. Intelligent tutoring systems, the pre-generative kind built around step-by-step problem solving, have produced effects approaching those of human tutors in some comparisons (VanLehn, 2011) and substantial median effects in meta-analysis, though smaller on standardised tests than on tests built for the system (Kulik & Fletcher, 2016). Adaptive software used in after-school centres in Delhi raised middle-school mathematics and Hindi scores substantially in under five months, including students who had been furthest behind (Muralidharan et al., 2019).

The generative AI evidence is newer. In a crossover trial at Harvard, undergraduate physics students learned more in less time from a carefully designed AI tutor than from an active-learning class (Kestin et al., 2025). In Nigeria, six weeks of after-school sessions with an AI tutor improved senior secondary students' results by about 0.3 standard deviations (De Simone et al., 2025). Tutors with an AI assistant helped more American students master topics (Wang et al., 2025), and a supervised AI tutor matched expert human tutors with UK students aged thirteen to fifteen (Google LearnLM Team & Eedi, 2025). Against that, unrestricted access to a general chatbot reduced learning in Turkish high schools (Bastani et al., 2025).

Three cautions follow. Most of these studies involve students aged thirteen or older. Most measure learning soon after the intervention, so retention is largely unknown. And context matters: a result in Lagos or Cambridge, Massachusetts, says something but not everything about Sydney. The reasonable conclusion is that a well-designed, supervised tutor can plausibly deliver effective practice for older primary students, that whether it can do so

conversationally for five- to eight-year-olds is unproven, and that the design should therefore restrict the role of conversational AI in the early years, as it now does.

28.2 Mastery learning

Mastery learning, as Bloom (1968) designed it, involves clear objectives, frequent formative checks, corrective teaching for students who have not yet met the standard and enrichment for those who have. A meta-analysis of mastery programmes found positive effects on achievement across many studies (Kulik et al., 1990). Slavin (1987) showed that effects are much smaller on standardised tests and in longer, larger implementations, and warned that the time needed for correctives can slow coverage. Guskey (2010) argues that the core elements, feedback, correctives and enrichment, are what matter, more than any particular programme. The PrimarAI hypothesis is that adaptive practice can reduce the cost of correctives that Slavin identified, letting mastery work at scale. That is plausible and untested. The specific thresholds in Section 11.6 have thinner support than the principle. The 80 per cent success rate comes from classroom observation studies (Rosenshine, 2012), the 85 per cent figure from a computational model (Wilson et al., 2019), and the 90 per cent gate from practice, not from trials. All three are starting values to be tuned in a pilot.

28.3 Explicit instruction and inquiry

This is where scholarly disagreement is real (Hmelo-Silver et al., 2007; Kirschner et al., 2006), as Section 8 discussed. The design takes a position: foundational knowledge is taught explicitly and practised (Castles et al., 2018; Rosenshine, 2012), and the projects in Create are structured and guided rather than left to discovery.

28.4 Experiential and project-based learning

A meta-analysis of project-based learning found a substantial positive average effect on achievement, with large variation by design, subject and duration (Chen & Yang, 2019). A cluster-randomised trial of a well-designed project-based science programme for Year 3 students improved their science achievement (Krajcik et al., 2023). Evidence from a range of designs links contact with nature to children's learning, attention and wellbeing (Kuo et al., 2019). Music is worth a specific note. A large multilevel meta-analysis found little evidence that music training improves children's general cognitive or academic skills once study quality is accounted for (Sala & Gobet, 2020). That does not weaken the case for music in this school so much as clarify it. Music is there because it is worth doing, not because it raises mathematics scores.

28.5 Agency and relationships

The evidence on motivation, self-regulation and teacher–student relationships reviewed in Part II (Roorda et al., 2011; Ryan & Deci, 2000; Zimmerman, 2002) supports the design's human structure. It does not prove that this particular structure is the best way to deliver it.

28.6 Success elsewhere is not proof here

None of this evidence shows that the PrimarAI School will work. Alpha's reported results, even if they were verified, would come from a selective, fee-charging American school. The Nigerian results come from secondary students in a low-resource setting. The strongest generalisation the evidence supports is that the components of the design each have reasonable support, and that the way they are combined is an open empirical question.

29 Objections and responses

Will students fall behind? This is the right first question. The earlier version of this paper answered it by claiming that NSW students are already "three to seven years behind the syllabus on average." That was wrong. Alpha reports that students transferring in from expensive private schools with straight As often assess years behind the grade on their report cards, which, if it holds up, says more about grading than about students. What the data show is a wide spread of achievement within every year level (Goss et al., 2016) and about one in three students below the Strong proficiency level in NAPLAN (The Educator, 2025). Mastery approaches are a reasonable response to that spread, but whether Augment keeps every student on track is unknown, so progress is checked each term against external tests, and students who are not progressing get more teacher time, not more screen time.

What about NESA, screens and teachers? The model is designed to be registrable, and the ARC week now sits close to NESA's time allocation advice in every learning area, including mathematics (Section 18). Four of the six daily sessions are screen-free (Section 24). And it needs more teachers, not fewer (Section 20.1).

What about privacy and Big Tech in classrooms? The charter commitments in Section 25 apply. The choice, though, is not really between Big Tech in the classroom and no AI. It is between purpose-built AI for students and students using ChatGPT on a parent's phone unsupervised, and we are already deep in the second scenario whether we like it or not.

Is this simply Alpha School with a hat? No. It is staffed by accredited teachers, not guides. It runs a different day, the ARC Day, with screen-free Reason sessions Alpha does not have, a rotation that gives the first session of the day to people as often as to software, and a taught

digital literacy strand Alpha does not offer. It is designed for inclusive admission, with a scholarship pathway, and for NESA alignment, and it does not stream students' screens. What it shares with Alpha is AI-tutored mastery practice and the idea that a large part of the school day can be given over to life. Alpha deserves the credit for proving that could be tried. Almost everything else is different, and the differences are the point. Table 1 in Section 3.1.3 sets out, feature by feature, what is taken, what is changed and what is left.

Could an existing school do this? Yes, and the ARC Day was redesigned so that one could. It runs on ordinary NSW bell times, so a school could trial the proposed day with a single class for a term without changing its building, its day or its other classes (Section 10.3).

AI is unpredictable, so how can anyone plan for it? Nobody can plan for it precisely, and pretending otherwise would be dishonest. What can be planned for is the capability set that will be useful across most plausible versions of the next twenty years. Knowledge, judgement, agency, relationships, creativity, embodied skill and ethical reasoning are useful across almost every plausible future, and valuable now whatever the future brings.

Is there a risk that this widens inequality? Yes, in both directions. If only wealthy families get AI tutoring, AI literacy and days like this, inequality widens. A fee-charging school can also widen it. The mitigations are a scholarship pathway from day one, open publication so any school can use any part of the design, partnerships with public schools and published results broken down by student background.

Is there evidence that AI tutoring works for young students? Not yet, at least not for conversational AI. That is why Augment in the early years is mostly teacher-led and why the youngest students use practice software rather than chatbots.

Are you simply swapping one prediction about the future for another? That was a fair criticism of earlier versions. The revised argument in Part II rests on what students deserve now, not on what the economy will reward later.

Is the tutor just drill software with a chat window? It could be, and that is the failure to design against. Three things are meant to prevent it: the Explain move, which asks for a reason before any step is checked; the rule that every content point is also taught by a teacher; and the unit lane, which ties practice to what the class is reading, measuring or making that week (Sections 11.6 and 18.4).

How do you know the syllabus is covered? Every step the tutor teaches is mapped to an outcome and a content point, every unit records which part of the day teaches, practises and

uses each content point, and the mastery record can be exported outcome by outcome for any student (Section 18). Appendix A shows this for two units.

30 Testing the model

30.1 A staged pilot

The model would be tested in three stages. In the first, design-based research in partner schools would trial parts of it that do not need a new timetable: digital literacy units, Reason practices and, with consent, short after-school trials of the tutor with Stage 2 and 3 students. The first tutor trial would be the pilot coaching course in Appendix A.3: two Department units, one Stage 1 class and one Stage 2 class, for a fortnight each. In the second, one class in an existing school would run the full ARC Day for a term, followed by a pilot of about fifty Stage 2 and 3 students over a full year. In the third, a school built on the model would open with evaluation built in from the first day, whether it is a new school or an existing one that adopts the ARC Day whole.

30.2 Evaluation design

An independent university partner would evaluate the pilot and the school against a pre-registered plan, using matched comparison groups of similar students in other schools, matched on prior achievement and background, together with interviews, observation and analysis of students' work. Results would be published whatever they show. Table 16 sets out the measures.

Table 16

Evaluation framework for the pilot and the school

Domain	Questions	Measures
Learning	Do students make at least as much progress in reading and mathematics as matched peers? Is the learning retained?	Standardised reading and mathematics tests each year, NAPLAN in Years 3 and 5, mastery records checked against these, delayed retention checks
Independent work	Can students read, write and solve problems without AI?	Blind-scored writing samples produced in Reason sessions; unaided problem-solving tasks
AI judgement	Can students direct, verify and question AI?	Performance tasks adapted from AI literacy frameworks
Wellbeing and engagement	Are students happy, safe and engaged?	Age-appropriate validated wellbeing measures, attendance, teacher and family observation

Domain	Questions	Measures
Teachers	Is the workload sustainable and the job rewarding?	Time-use diaries, surveys, retention
Equity	Does the model work for every group of students?	All outcomes reported by disability, background, language and scholarship status
Cost	What does it cost, and could it be afforded elsewhere?	Cost per student, staffing ratios, comparison with similar schools
Family confidence	Do families trust and value the school?	Surveys, re-enrolment, complaints
Unintended consequences	What is going wrong that nobody planned for?	Gap between performance with and without AI, screen-time totals, privacy and safeguarding incidents

31 What would count against the model

A design that cannot be wrong cannot be tested. These are the results that would require PrimarAI to change, stated before any data exist.

1. If, after two years, students' growth in reading or mathematics is meaningfully below that of matched peers, Augment is restructured, whether with more teacher-led instruction, more time or a different tool.
2. If students' writing produced without AI is weaker than that of matched peers, the balance between Reason and Create changes.
3. If learning in the first session of the day shows no advantage over learning later in the day, the rotation is dropped in favour of a simpler fixed timetable.
4. If students perform well with AI support and poorly without it, the signature of the performance paradox, the tutor's design and role are reduced until the gap closes.
5. If wellbeing, engagement or attendance decline, the programme is reviewed with families and students.
6. If teacher workload rises or teachers leave, the staffing model is redesigned.
7. If any group of students does worse than its matched peers, that is treated as a design failure, and the model is not expanded until it is fixed.
8. If the school cannot be funded without abandoning small human groups, the design changes publicly rather than quietly.
9. If digital literacy does not improve students' judgement about AI compared with peers, the strand is rethought.
10. Any serious privacy breach suspends the tool involved pending independent review.

In every case the results and the changes will be published.

32 Implementation and the conditions for success

Turning the design into a school is a staged project. It starts with an entity, a board and university and school partners, followed by a full financial model, the curriculum Blueprint and NESA registration documentation. Then come the technology build or procurement, with an independent privacy impact assessment, then teacher recruitment and the teachers' programme, the pilot, premises and planning approvals, which can take years in NSW and need to start early, and finally registration and launch. Three coordinated programmes run throughout: a teachers' programme, a student and partner programme, and a technology programme with Australian hosting and transparent use.

The model will succeed only if a handful of conditions hold: enough excellent teachers who want to work this way; a financial model that sustains small human groups; a tutor that meets the design standard, which means Socratic, supervised, privacy-preserving and aligned to the NSW syllabuses; families who trust the school; community partners who stay; and an evaluation honest enough to change the design. None of those is guaranteed, but all of them can be worked on in public.

Epilogue: An invitation to build the next school

So, can you be mean to a chatbot? I still do not know the answer about the chatbot. I am fairly sure about the student. A student who learns to be cruel to something that talks back is practising cruelty, and a student who learns to ask the question at all is practising something better. That is what this school is for: not to settle the questions AI raises, but to make sure students grow up able to ask them, argue about them and act well while the answers are still uncertain.

The proposal in this paper is simple to state. AI gives us an opportunity to reconsider the architecture of schooling itself, and to return time, attention and experience to students and teachers. The ARC Day is one way of doing that, and the latest of several. It may well be wrong in its details, and the evaluation in Part VI is designed to find out where.

I am one teacher in the Eastern Suburbs of Sydney with a kid in primary school and a calculator. I cannot build this on my own, and there is no point pretending otherwise. The Castlereagh Statement is the closest thing Australia has to a national consensus on AI in education. Its seventy contributors include people I respect, among them Danny Liu, Jason Lodge, Leon Furze, Matt Bower and Simon Buckingham Shum, alongside university leaders, school principals and public servants. The Statement is excellent, but it has a hole in the shape of primary school, and it leaves the far horizon, where schooling is actually restructured, as an exercise for the reader. PrimarAI is my attempt to work on that horizon in public, showing my working as I go. Over the coming months I will publish the full Blueprint in stages, covering pedagogy, timetable, syllabus mapping, staffing, costs, the NESA pathway and the evaluation plan. It will be free to read, free to argue with and free to copy, and it will be aligned explicitly with Castlereagh's principles and horizons so that the Australian conversation about Kindergarten to Year 6 has a worked example on the table.

If you are a teacher who is tired, a principal who is curious, a parent who is worried, a researcher who would like to test something real, a board chair with a half-empty enrolment, an investor interested in what comes next, a Castlereagh contributor who quietly wonders what any of this means for an eight-year-old, or simply someone with a strong opinion and a reason to engage, I want you in the conversation. The work is bigger than any one person can do alone, and the design only really gets interesting when other people start arguing with it.

> *The school model we have is wrong for the era. I would like your help to build the next one, with teachers rather than around them, and with students, who, believe me, already know.*

Dane Smale, Sydney, November 2026

References

Acemoglu, D. (2025). The simple macroeconomics of AI. *Economic Policy*, *40*(121), 13–58. https://doi.org/10.1093/epolic/eiae042

Aldeman, C. (2026, January 8). What public schools and parents can learn from a $40,000-a-year private school. *The 74*. https://www.the74million.org/article/what-public-schools-and-parents-can-learn-from-a-40000-a-year-private-school/

Alexander, R. (2020). *A dialogic teaching companion*. Routledge.

Alpha School. (n.d.-a). *Frequently asked questions*. Retrieved November 14, 2026, from https://alpha.school/faq/

Alpha School. (n.d.-b). *The two-hour school day: How AI tutors are redefining learning efficiency*. Retrieved November 14, 2026, from https://alpha.school/?p=8134

Alpha School. (n.d.-c). *What does Alpha School's two-hour academic day look like?* Retrieved November 14, 2026, from https://alpha.school/resources/how-alpha-two-hour-school-day-works/

Alpha School. (n.d.-d). *How does Alpha School use personalized learning to support student mastery and self-directed learning?* Retrieved November 14, 2026, from https://alpha.school/resources/how-does-alpha-school-use-personalized-learning-to-support-student-mastery-and-s/

Alpha School. (n.d.-e). *Life skills workshops*. Retrieved November 14, 2026, from https://alpha.school/lifeskills-workshops

Alpha School. (n.d.-f). *The program*. Retrieved November 14, 2026, from https://alpha.school/the-program/

Alpha School. (n.d.-g). *Learning science team*. Retrieved November 14, 2026, from https://alpha.school/learning-science/

Alpha School. (2026). *SY26–27 session 1: Guides*. https://alpha.school/sy26-27-session-1-guides/

Alvarado, R. (2023). AI as an epistemic technology. *Science and Engineering Ethics*, *29*(5), Article 32. https://doi.org/10.1007/s11948-023-00451-3

Amodei, D. (2024, October). *Machines of loving grace: How AI could transform the world for the better*. https://darioamodei.com/machines-of-loving-grace

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, Article 87. https://doi.org/10.3389/feduc.2019.00087

Andries, V., & Robertson, J. (2023). "Alexa doesn't have that many feelings": Children's understanding of AI through interactions with smart speakers in their homes. *Computers and Education: Artificial Intelligence*, *5*, Article 100176. https://doi.org/10.1016/j.caeai.2023.100176

Anthropic. (2026, May 14). [Announcement of a US$200 million partnership with the Gates Foundation]. https://www.anthropic.com/news/gates-foundation-partnership

Austin Scholar. (2025). *Austin Scholar #177: The Alpha app stack*. Substack. https://austinscholar.substack.com/p/austin-scholar-177-the-alpha-app

Australian Education Regulations 2023 (Cth).

Australian Government Department of Education. (2025). *China's approach to AI education in schools (Year 1–12)*. https://www.education.gov.au/download/19495/chinas-approach-ai-education-primary-and-secondary-schools/41945/document/pdf

Australian Government Department of Health. (2021). *Australian 24-hour movement guidelines for children (5 to 12 years) and young people (13 to 17 years): An integration of physical activity, sedentary behaviour and sleep*. https://www.health.gov.au/sites/default/files/documents/2021/03/australian-24-hour-movement-guidelines-for-children-5-to-12-years-and-young-people-13-to-17-years-an-integration-of-physical-activity-sedentary-behaviour-and-sleep.pdf

Bashir, D. (Host). (2024, July 4). Dan Hart and Michelle Michael: Bringing AI to students in New South Wales (No. 129) [Audio podcast episode]. In *The Gradient: Perspectives on AI*. https://thegradientpub.substack.com/p/dan-hart-michelle-michael-ai-education-nsw

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, *122*(26), Article e2422633122. https://doi.org/10.1073/pnas.2422633122

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. Columbia University Press.

Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, *9*(4), 475–479. https://doi.org/10.1016/j.jarmac.2020.09.003

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment*, *1*(2), 1–12.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*(6), 4–16. https://doi.org/10.3102/0013189X013006004

Bowen, D., & Fleming, R. (Hosts). (2026, February 22). NSW EduChat and the largest deployment of generative AI to students [Audio podcast episode with D. Hart]. In *AI in Education*.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies: "Engines of growth"? *Journal of Econometrics*, *65*(1), 83–108. https://doi.org/10.1016/0304-4076(94)01598-T

Brynjolfsson, E., Chandar, B., & Chen, R. (2025). *Canaries in the coal mine? Six facts about the recent employment effects of artificial intelligence* [Working paper]. Stanford Digital Economy Lab. https://digitaleconomy.stanford.edu/publications/canaries-in-the-coal-mine

Carney, M., Webster, B., Alvarado, I., Phillips, K., Howell, N., Griffith, J., Jongejan, J., Pitaru, A., & Chen, A. (2020). Teachable Machine: Approachable web-based tool for exploring machine learning classification. In *Extended abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–8). Association for Computing Machinery. https://doi.org/10.1145/3334480.3382839

CAST. (2024). *Universal design for learning guidelines version 3.0*. https://udlguidelines.cast.org

Castlereagh Summit. (2026). *The Castlereagh Statement: A cross-sector call to action on Australian education and training in the age of AI*. https://castlereagh.ai/wp-content/uploads/2026/04/The-Castlereagh-Statement-v1.00.pdf

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Chatterji, A., Cunningham, T., Deming, D. J., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025). *How people use ChatGPT* (NBER Working Paper No. 34255). National Bureau of Economic Research. https://doi.org/10.3386/w34255

Chen, B. (2025). *Beyond tools: Generative AI as epistemic infrastructure in education* (Version 2) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2504.06928

Chen, C.-H., & Yang, Y.-C. (2019). Revisiting the effects of project-based learning on students' academic achievement: A meta-analysis investigating moderators. *Educational Research Review*, *26*, 71–81. https://doi.org/10.1016/j.edurev.2018.11.001

Children's Guardian Act 2019 (NSW).

Cobbold, T. (2025, November 20). Overworked, overburdened, and burning out: Australian teachers' workloads among the worst in OECD. *Pearls and Irritations*. https://www.pearlsandirritations.com/post/2025/11/oecd-survey-shows-australian-teachers-have-a-very-heavy-workload/

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. Teachers College Press.

David, P. A. (1990). The dynamo and the computer: An historical perspective on the modern productivity paradox. *American Economic Review*, *80*(2), 355–361.

De Bruyckere, P. (2026, September 1). Do we know the alpha and omega of Alpha School? *From Experience to Meaning*. https://theeconomyofmeaning.com/2026/09/01/do-we-know-the-alpha-and-omega-of-alpha-school/

De Simone, M., Tiberti, F., Barron Rodriguez, M., Manolio, F., Mosuro, W., & Dikoru, E. J. (2025). *From chalkboards to chatbots: Evaluating the impact of generative AI on learning outcomes in Nigeria* (Policy Research Working Paper No. 11125). World Bank. https://doi.org/10.1596/1813-9450-11125

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627–668. https://doi.org/10.1037/0033-2909.125.6.627

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, *25*, 23–38. https://doi.org/10.1016/j.edurev.2018.09.003

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, *227*, Article 105224. https://doi.org/10.1016/j.compedu.2024.105224

Disability Standards for Education 2005 (Cth).

Duggan, S. (2026, March 13). Surface-level mastery, cognitive atrophy a risk of schools' unchecked AI use. *EducationHQ*. https://educationhq.com/news/surface-level-mastery-cognitive-atrophy-a-risk-of-schools-unchecked-ai-use-207204/

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Education Endowment Foundation. (2018). *GraphoGame Rime: Evaluation report and executive summary*. https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/graphogame-rime

Feathers, T. (2025, October 27). Parents fell in love with Alpha School's promise. Then they wanted out. *WIRED*. https://www.wired.com/story/ai-teacher-inside-alpha-school/

Maiberg, E. (2026). "Students are being treated like guinea pigs": Inside an AI-powered private school. *404 Media*. https://404media.co/students-are-being-treated-like-guinea-pigs-inside-an-ai-powered-private-school

NextGen Academy. (n.d.). *NextGen Academy*. Retrieved November 14, 2026, from https://nextgenacademy.school/

Nova Academy. (n.d.). *Nova Academy*. Retrieved November 14, 2026, from https://novaacademy.school/

NSW Department of Education. (2025). *Mathematics Stage 2: Unit 20* [Teaching unit]. State of New South Wales.

NSW Department of Education. (2026). *English Stage 1: Unit 20, Colour Me* [Teaching unit]. State of New South Wales.

The Educator. (2025, July 30). NAPLAN national results released. *The Educator K/12*. https://www.theeducatoronline.com/k12/news/naplan-national-results-released/287677

Ellington, A. J. (2003). A meta-analysis of the effects of calculators on students' achievement and attitude levels in precollege mathematics classes. *Journal for Research in Mathematics Education*, *34*(5), 433–463. https://doi.org/10.2307/30034795

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, *384*(6702), 1306–1308. https://doi.org/10.1126/science.adj0998

European Commission & OECD. (2025). *Empowering learners for the age of AI: An AI literacy framework for primary and secondary education* (Review draft). https://ailiteracyframework.org

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, *56*(2), 489–530. https://doi.org/10.1111/bjet.13544

Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A comparison of children's reading on paper versus screen: A meta-analysis. *Review of Educational Research*, *91*(4), 483–517. https://doi.org/10.3102/0034654321998074

Generation AI: Fears of "social divide" unless all children learn computing skills. (2026, January 5). *The Guardian*. https://www.theguardian.com/education/2026/jan/05/generation-ai-fears-of-social-divide-unless-all-children-learn-computing-skills

Google LearnLM Team & Eedi. (2025). *AI tutoring can safely and effectively support students: An exploratory RCT in UK classrooms* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2512.23633

Goss, P., Sonnemann, J., Chisholm, C., & Nelson, L. (2016). *Widening gaps: What NAPLAN tells us about student progress* (Report No. 2016-3). Grattan Institute. https://grattan.edu.au/wp-content/uploads/2016/03/937-Widening-gaps.pdf

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, *104*(4), 879–896. https://doi.org/10.1037/a0029185

Guskey, T. R. (2010). Lessons of mastery learning. *Educational Leadership*, *68*(2), 52–57.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*(2), 625–638. https://doi.org/10.1111/1467-8624.00301

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Health Records and Information Privacy Act 2002 (NSW).

Hill, A. J., & Jones, D. B. (2018). A teacher who knows me: The academic benefits of repeat student-teacher matches. *Economics of Education Review*, *64*, 1–12. https://doi.org/10.1016/j.econedurev.2018.03.004

Hills, L., & Henkel, O. (Hosts). (2025, August 14). Assessment in education: To AI or not to AI? (Season 2, Episode 15) [Audio podcast episode with D. Wiliam]. In *Ed-Technical*. https://boldscience.org/assessment-in-education-to-ai-or-not-to-ai/

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368

Humlum, A., & Vestergaard, E. (2025). *Large language models, small labor market effects* (Working Paper No. 2025-56). Becker Friedman Institute for Economics, University of Chicago. https://bfi.uchicago.edu/wp-content/uploads/2025/04/BFI_WP_2025-56-3.pdf

Hunter, J., Sonnemann, J., & Joiner, R. (2022). *Making time for great teaching: How better government policy can help*. Grattan Institute.

Jobs and Skills Australia. (2025). *Our Gen AI transition: Implications for work and skills*. https://www.jobsandskills.gov.au/publications/generative-ai-capacity-study-report

Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., Ruckert, J. H., & Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental Psychology*, *48*(2), 303–314. https://doi.org/10.1037/a0027033

Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist*, *51*(2), 289–299. https://doi.org/10.1080/00461520.2016.1155457

Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, *15*, Article 17458. https://doi.org/10.1038/s41598-025-97652-6

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2506.08872

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Krajcik, J., Schneider, B., Miller, E. A., Chen, I.-C., Bradford, L., Baker, Q., Bartz, K., Miller, C., Li, T., Codere, S., & Peek-Brown, D. (2023). Assessing the effect of project-based learning on science learning in elementary schools. *American Educational Research Journal*, *60*(1), 70–102. https://doi.org/10.3102/00028312221129247

Kulik, C.-L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, *60*(2), 265–299. https://doi.org/10.3102/00346543060002265

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, *86*(1), 42–78. https://doi.org/10.3102/0034654315581420

Kuo, M., Barnes, M., & Jordan, C. (2019). Do experiences with nature promote learning? Converging evidence of a cause-and-effect relationship. *Frontiers in Psychology*, *10*, Article 305. https://doi.org/10.3389/fpsyg.2019.00305

Kurian, N. (2025). 'No, Alexa, no!': Designing child-safe AI and protecting children from the risks of the 'empathy gap' in large language models. *Learning, Media and Technology*, *50*(4), 621–634. https://doi.org/10.1080/17439884.2024.2367052

Kwa, T., West, B., Becker, J., Deng, A., Garcia, K., Hasin, M., Jawhar, S., Kinniment, M., Rush, N., Von Arx, S., Bloom, R., Broadley, T., Du, H., Goodrich, B., Jurkovic, N., Miles, L. H., Nix, S., Lin, T., Painter, C., . . . Chan, L. (2026). *Measuring AI ability to complete long software tasks* (Version 3) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2503.14499

Lehmann, M., Cornelius, P. B., & Sting, F. J. (2024). *AI meets the classroom: When do large language models harm learning?* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2409.09047

Lodge, J. M., Yang, S., Furze, L., & Dawson, P. (2023). It's not like a calculator, so what is the relationship between learners and generative artificial intelligence? *Learning: Research and Practice*, *9*(2), 117–124. https://doi.org/10.1080/23735082.2023.2261106

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727

Masters, G. (2020). *Nurturing wonder and igniting passion: Designs for a new school curriculum* [NSW Curriculum Review final report]. NSW Education Standards Authority.

McCain, M., Linthicum, R., Lubinski, C., Tamkin, A., Huang, S., Stern, M., Handa, K., Durmus, E., Neylon, T., Ritchie, S., Jagadish, K., Maheshwary, P., Heck, S., Sanderford, A., & Ganguli, D. (2025, June 27). *How people use Claude for support, advice, and companionship*. Anthropic. https://www.anthropic.com/research/how-people-use-claude-for-support-advice-and-companionship

Meyer, D. (2026, August 12). Does Alpha School work for regular kids? *Mathworlds*. https://danmeyer.substack.com/p/does-alpha-school-work-for-regular

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Miao, F., & Shiohira, K. (2024). *AI competency framework for students*. UNESCO. https://doi.org/10.54675/JKJB9835

Ministry of Education and Research. (2025, February 25). *Estonia announces a groundbreaking national initiative: AI Leap programme to bring AI tools to all schools* [Press release]. Republic of Estonia. https://hm.ee/en/news/estonia-announces-groundbreaking-national-initiative-ai-leap-programme-bring-ai-tools-all

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review*, *109*(4), 1426–1460. https://doi.org/10.1257/aer.20171112

Nickow, A., Oreopoulos, P., & Quan, V. (2024). The promise of tutoring for PreK–12 learning: A systematic review and meta-analysis of the experimental evidence. *American Educational Research Journal*, *61*(1), 74–107. https://doi.org/10.3102/00028312231208687

Noles, C. (2026, April 29). What Greg Brockman thinks comes after ChatGPT. *The Neuron*. https://www.theneuron.ai/explainer-articles/what-greg-brockman-thinks-comes-after-chatgpt/

NSW Department of Education. (2024). *Sport and physical activity* (PD-2004-0034-11-V01.2.1) [Policy]. https://education.nsw.gov.au/policy-library/policies/pd-2004-0034-11

NSW Education Standards Authority. (2022a). *English K–10 syllabus*. NSW Government. https://curriculum.nsw.edu.au

NSW Education Standards Authority. (2022b). *Mathematics K–10 syllabus*. NSW Government. https://curriculum.nsw.edu.au

NSW Education Standards Authority. (2026). *Curriculum: School registration requirements* [Fact sheet]. https://www.nsw.gov.au/sites/default/files/2026-09/fact-sheet-curriculum-august-2026.pdf

NSW Education Standards Authority. (n.d.-a). *Curriculum reform timeline*. Retrieved October 20, 2026, from https://www.nsw.gov.au/education-and-training/nesa/about/strategies-and-reforms/curriculum-reform/timeline

NSW Education Standards Authority. (n.d.-b). *K–6 time allocation advice*. NSW Curriculum. Retrieved October 24, 2026, from https://curriculum.nsw.edu.au/use-the-curriculum/planning-supporting-student-learning/k-6-time-allocation-advice

NSW Government. (2025, September 23). *NSW to roll out purpose-built AI education tool to all public school students from Year 5* [Media release]. https://www.nsw.gov.au/ministerial-releases/nsw-to-roll-out-purpose-built-ai-education-tool-to-all-public-school-students-from-year-5

Nussbaum, M. C. (2010). *Not for profit: Why democracy needs the humanities*. Princeton University Press.

OECD. (2025a). *Education for human flourishing: A conceptual framework*. OECD Publishing. https://doi.org/10.1787/73d7cb96-en

OECD. (2025b). *Results from TALIS 2024: The state of teaching*. OECD Publishing. https://doi.org/10.1787/90df6235-en

OECD. (2026). *OECD digital education outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. https://doi.org/10.1787/062a7394-en

Office of the Australian Information Commissioner. (2026). *Draft Children's Online Privacy Code: Consultation for industry, civil society, academia and other interested stakeholders*. https://www.oaic.gov.au/engage-with-us/consultations/draft-childrens-online-privacy-code-consultation-for-industry,-civil-society,-academia

Online Safety Amendment (Social Media Minimum Age) Act 2024 (Cth).

OpenAI. (2025, April 29). *Sycophancy in GPT-4o: What happened and what we're doing about it*. https://openai.com/index/sycophancy-in-gpt-4o/

Plato. (1995). *Phaedrus* (A. Nehamas & P. Woodruff, Trans.). Hackett. (Original work published ca. 370 B.C.E.)

Pope, N. G. (2016). How the time of day affects productivity: Evidence from school schedules. *The Review of Economics and Statistics*, *98*(1), 1–11. https://doi.org/10.1162/REST_a_00525

Privacy Act 1988 (Cth).

Privacy and Other Legislation Amendment Act 2024 (Cth).

Rasheed, A. (2025, August 23). UAE schools to introduce AI curriculum in new academic year. *Gulf News*. https://gulfnews.com/uae/education/uae-schools-to-introduce-ai-curriculum-in-new-academic-year-1.500242831

Reich, J. (2020). *Failure to disrupt: Why technology alone can't transform education*. Harvard University Press.

Reich, J., & Dukes, J. (Hosts). (2025). *The homework machine* [Audio podcast series]. MIT Teaching Systems Lab. https://cmsw.mit.edu/homework-machine-justin-reich-jesse-dukes/

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Robb, M. B., & Mann, S. (2025). *Talk, trust, and trade-offs: How and why teens use AI companions*. Common Sense Media. https://www.commonsensemedia.org/research/talk-trust-and-trade-offs-how-and-why-teens-use-ai-companions

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, *81*(4), 493–529. https://doi.org/10.3102/0034654311421793

Rosenshine, B. (2012). Principles of instruction: Research-based strategies that all teachers should know. *American Educator*, *36*(1), 12–19, 39.

Royal Life Saving Society – Australia. (2020, August 10). *New National Swimming and Water Safety Framework*. https://www.royallifesaving.com.au/about/news-and-updates/news/2020/aug/framework

Runaway Sydney Half Marathon. (n.d.). *Half marathon*. Retrieved October 20, 2026, from https://runawaysydneyhalf.com.au/race-options/half-marathon/

Ryan, A. M., Pintrich, P. R., & Midgley, C. (2001). Avoiding seeking help in the classroom: Who and why? *Educational Psychology Review*, *13*(2), 93–114. https://doi.org/10.1023/A:1009013420053

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Sala, G., & Gobet, F. (2020). Cognitive and academic benefits of music training with children: A multilevel meta-analysis. *Memory & Cognition*, *48*(8), 1429–1441. https://doi.org/10.3758/s13421-020-01060-2

Sands, A. (2026, August 2). AI-powered Alpha Schools expands to 50 campuses. *Axios*. https://www.axios.com/2026/08/02/alpha-schools-ai-expansion-50-campuses

Santangelo, T., & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review*, *28*(2), 225–265. https://doi.org/10.1007/s10648-015-9335-1

Selitskiy, S., & Inoue, C. (2025). *Education paradigm shift to maintain human competitive advantage over AI* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2510.23436

Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R., Cheng, N., Durmus, E., Hatfield-Dodds, Z., Johnston, S. R., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., Zhang, M., & Perez, E. (2023). *Towards understanding sycophancy in language models* (Version 4) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2310.13548

Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences*, *113*(10), 2621–2624. https://doi.org/10.1073/pnas.1516947113

Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research*, *91*(5), 761–798. https://doi.org/10.3102/00346543211019105

Slavin, R. E. (1987). Mastery learning reconsidered. *Review of Educational Research*, *57*(2), 175–213. https://doi.org/10.3102/00346543057002175

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, *160*, Article 108386. https://doi.org/10.1016/j.chb.2024.108386

Stern, J. (2026). *I am not a robot: My year using AI to do (almost) everything and replace (almost) everyone*. Harper.

Swisher, K., & Galloway, S. (Hosts). (2026, April 30). One year inside AI tools [Audio podcast episode with J. Stern]. In *Pivot*. Vox Media. https://podcasts.voxmedia.com/show/pivot

Thian, S. Y. (2026). Estonia exercises sovereignty in its ministry-led AI movement for schools. *GovInsider*. https://govinsider.asia/intl-en/article/estonia-exercises-sovereignty-in-its-moe-led-ai-movement-for-schools

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What should every child know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(1), 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

University of Technology Sydney. (2026, July 16). *When AI does the homework, are students still learning?* https://www.uts.edu.au/vi/news/2026/07/when-ai-does-the-homework-are-students-still-learning

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis. *Review of Educational Research*, *65*(4), 319–381. https://doi.org/10.3102/00346543065004319

von Hippel, P. T. (2024). Two-sigma tutoring: Separating science fiction from science fact. *Education Next*, *24*(2). https://www.educationnext.org/two-sigma-tutoring-separating-science-fiction-from-science-fact/

Wang, R. E., Ribeiro, A. T., Robinson, C. D., Loeb, S., & Demszky, D. (2025). *Tutor CoPilot: A human-AI approach for scaling real-time expertise* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2410.03017

Watters, A. (2021). *Teaching machines: The history of personalized learning*. MIT Press.

Waypoint Academy. (n.d.). *Waypoint Academy*. Retrieved November 14, 2026, from https://waypointacademy.school/

Weidlich, J., Gašević, D., Drachsler, H., & Kirschner, P. (2025). ChatGPT in education: An effect in search of a cause. *Journal of Computer Assisted Learning*, *41*(5), Article e70105. https://doi.org/10.1111/jcal.70105

Wheatley, M. (2026, May 11). Thinking Machines drops a new, highly responsive model designed for humanlike interactions in real time. *SiliconANGLE*. https://siliconangle.com/2026/05/11/thinking-machines-drops-new-highly-responsive-model-designed-humanlike-interactions-real-time/

Willingham, D. T. (2007). Critical thinking: Why is it so hard to teach? *American Educator*, *31*(2), 8–19.

Wilson, R. C., Shenhav, A., Straccia, M., & Cohen, J. D. (2019). The eighty five percent rule for optimal learning. *Nature Communications, 10*, Article 4646. https://doi.org/10.1038/s41467-019-12552-4

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, *10*, Article 3087. https://doi.org/10.3389/fpsyg.2019.03087

Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., Hessert, W. T., Williams, M. E., & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, *143*(2), 804–824. https://doi.org/10.1037/a0033906

Young, M., & Muller, J. (2013). On the powers of powerful knowledge. *Review of Education*, *1*(3), 229–250. https://doi.org/10.1002/rev3.3017

Zao-Sanders, M. (2025, April 9). How people are really using gen AI in 2025. *Harvard Business Review*. https://hbr.org/2025/04/how-people-are-really-using-gen-ai-in-2025

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Appendix A: Two NSW units taught the ARC way

This appendix takes two current units published by the NSW Department of Education and shows where each lesson would sit in the proposed ARC Day, with one Augment session and one Reason session planned in full for each. The outcomes, content points and texts are the Department's. The division across the day and the session plans are mine. The original units are © State of New South Wales (Department of Education) and are licensed under CC BY 4.0 (NSW Department of Education, 2025, 2026).

A.1 English, Stage 1: Unit 20, Colour Me

The unit runs for two weeks in ten 40-minute lessons. Students make inferences from words and illustrations, study figurative language and wordplay in a poem and a picture book, and then plan, write, edit and illustrate a poem of their own. The mentor text is *Colour Me* by Ezekiel Kwaymullina, illustrated by Moira Court, and the supporting text is the poem "A Leaf Mountain" by Sally Murphy. The outcomes are EN1-OLC-01, EN1-VOCAB-01, EN1-RECOM-01, EN1-CWT-01 and EN1-UARL-01.

Because this is a Stage 1 unit, Section 11.2 applies: teachers lead, there is no chat, and screen time stays within about 45 minutes a day. The unit is built around two whole texts, a great deal of talk and one piece of writing, which is the work of Reason. Its ten lessons map one to one onto the ten Reason English sessions of a fortnight. The four English Augment sessions each week keep their standing job, which is phonics, spelling, handwriting and fluency in small groups with a teacher, and give one rotation station to the unit's skill-based content: inference, action verbs, noun groups and words in context.

Table A1

English Stage 1 Unit 20 across a proposed ARC fortnight

Day	Reason: the class, screens off	Augment: the station for this unit	Create
Week 1 Monday	Lesson 1: making inferences from the words and images of the poem	Inference pairs: two sentences read aloud, choose the likely inference	
Tuesday	Lesson 2: wordplay, repetition and rhythm; write two poem lines	Inference pairs at each student's level, now with a picture	Drama and dance: perform the action lines with movement

Day	Reason: the class, screens off	Augment: the station for this unit	Create
Wednesday	Lesson 3: patterns and figurative language in the mentor text	Action verbs: choose the verb that fits a picture	
Thursday (long session)	Lesson 4: modifying words; write a sentence that compares a colour to a noun group	Noun groups: add a modifying word to a noun	Visual arts: colour mixing and collage in the manner of the illustrations
Friday	Lesson 5: how illustrations add meaning, run as Stage 1's Friday argument: do the words or the pictures tell you more?	No English Augment on Friday	
Week 2 Monday	Lesson 6: words in different contexts; a four-line rainbow poem	Noun groups; delayed check on inference	
Tuesday	Lesson 7: stating a preference between the two texts, with reasons	Words in context: one word, two meanings	
Wednesday	Lesson 8: planning a colour poem	Words in context; adjectives that fit a noun	
Thursday (long session)	Lesson 9: writing the poem by hand	Mixed review and delayed checks	Visual arts: begin the illustration
Friday	Lesson 10: editing from teacher feedback, illustrating and sharing	No English Augment on Friday	Community and showcase: poems read aloud and displayed

Note. Days and sessions follow the example ARC week in Table 4. Lesson titles are paraphrased from the Department's unit.

Session plan 1: Augment, Stage 1 English, Week 1 Tuesday, 1:15 to 2:00

One teacher works with about twelve students, in three groups of four.

1. **Launch, 3 minutes.** The twelve sit together. The teacher reads two sentences aloud and students show A or B on their fingers, which is the routine the Department's lesson opens with.
2. Three rotations of 12 minutes, with a minute to change over.
   - **Teacher table.** The day's phonics and spelling focus from the school's sequence, such as a vowel digraph (EN1-PHOKW-01, EN1-SPELL-01), ending with a dictated sentence written by hand.
   - **Tablet station, with headphones.** About six minutes of inference pairs for this unit, then about six minutes at the student's own edge in phonics or spelling.

   - **Hands-on station, no screen.** Students match sentence strips to the picture they describe and tell a partner what they infer, using "I think … because …".
3. **Close, 3 minutes.** Each student tells a partner one "I think … because …" sentence about a new picture. The teacher notes who cannot yet give a reason.

Each student spends about 12 minutes of this session on a screen.

What the tablet station does. Items come in four levels: a picture with one sentence; two sentences with two choices, as in the Department's lesson; three sentences with three choices, one of which is true but not supported by the text; and a short paragraph. Everything is read aloud. After choosing, the student taps the words that gave the clue, which is the Explain move in a form a six-year-old can manage. After a first wrong answer the sentences are read again with one clue highlighted. After a second, the item goes on the teacher's list and the next item is easier. Nine in ten at the third level across two days marks the step ready, a delayed check follows the next week, and the spoken reason is always checked by a teacher.

Sample items, written for this paper:

- "Ava put on her hat and rubbed sunscreen on her nose. She picked up her towel." Where is Ava going? A: to the beach. B: to bed.
- "The puppy's bowl was upside down. There was water all over the kitchen floor." What happened? A: it rained inside. B: the puppy knocked the bowl over.
- "Tom looked at the clock. He grabbed his bag and ran to the bus stop." What can we infer? A: Tom is late. B: Tom likes buses. C: Tom has a bag.

Session plan 2: Reason, Stage 1 English, Lesson 1, Monday 11:15 to 12:00

This session follows the Department's lesson almost as written, because the lesson is already what Reason is for.

1. **Beginning, 5 minutes.** Inference pairs on mini whiteboards to bring back what students know.
2. **Core, 30 minutes.** The teacher reads the poem aloud from a paper display, thinks aloud to connect meaning across lines, and links the words to the illustration. Students turn and talk with "The words make me imagine … because …", the class builds an anchor chart of words that help a reader imagine, and each student records two examples with a sketch.

3. **Writing, 5 minutes.** Every student writes one sentence by hand to explain a sketch. In the Department's lesson this is the extension task. Here it is for everyone, because the session is five minutes longer.
4. **Concluding, 5 minutes.** Students share a sketch with a partner and say which words led to it.

What travels between the two sessions. From Reason to Augment: the teacher's note of who could not yet make an inference from two sentences. Those students start at the picture level on the tablet and go to the teacher table first. From Augment to Reason: the list of students who chose the true-but-unsupported answer, which tells the teacher what to model on Wednesday.

Evidence for the unit. The mastery record shows the inference, noun group and words-in-context steps. The teacher's observations use the "what to look for" questions in each Department lesson. The portfolio holds the sketches, the rainbow poem, and the first and edited drafts of the final poem.

A.2 Mathematics, Stage 2: Unit 20

The unit has eight lessons of 60 to 75 minutes and eight short number sense activities. It covers place value, additive and multiplicative relations and grid maps. The outcomes are MA2-RN-01, MA2-AR-01, MA2-AR-02, MA2-MR-01, MA2-MR-02, MA2-GM-01 and MAO-WM-01.

Each Department lesson has a teaching part, in which the teacher brings back prior knowledge, teaches explicitly and leads a discussion of strategies, and a practice part, usually a game or task. In the proposed ARC week the teaching part goes to Reason, three sessions a week, and the practice part goes to Augment, four sessions a week. There each student practises the same content at their own level in the unit lane and spends the rest of the session at their own edge (Section 18.4). The number sense activities, which the Department designed to activate prior knowledge and to loop back for spaced practice, become the Retrieve move that opens every Augment session.

Table A2

Mathematics Stage 2 Unit 20 across a proposed ARC fortnight

Day	Augment: the unit lane	Reason: the class, screens off	Create
Week 1 Monday	Retrieve with number sense activities 1 and 2 (mental addition); then the edge lane	Lesson 3, part 1: partitioning four-digit numbers in non-standard ways, with materials	Making and enterprise: the garden team costs seedlings for the parent fundraiser
Tuesday	No mathematics Augment	Lesson 3, part 2: what happens to each digit when a number is 10, 100 or 1000 times as large. Lesson 2 begins: dollars and cents	
Wednesday	Place value: partition, rename, and make numbers 10, 100 and 1000 times as large	No mathematics Reason	
Thursday	Money: equivalent amounts, totals and change	Mathematics outside. Lesson 1: grid maps. Students map the school grounds on a grid and write grid references	
Friday	Grid references; first checks on the place value steps	No mathematics Reason	Sport: a treasure hunt using each other's grid maps
Week 2 Monday	Delayed checks on place value; number sense activities 3 and 4	Lesson 4: compensation, levelling and constant difference; the card game Closest to 100 in pairs	Making and enterprise: a price list and a float for the seedling table at the fundraiser
Tuesday	No mathematics Augment	Lesson 5: the equals sign as balance. Lesson 6: the link between addition and subtraction	
Wednesday	Coached practice on the three strategies, then choose one and explain (session plan 3)	No mathematics Reason	
Thursday	Equality and missing values; using addition to solve subtraction	Mathematics outside. Lesson 7: arrays in the garden, four rows of six seedlings and six rows of four	
Friday	Fact families for 2, 4, 5 and 10 with spaced retrieval; checks on the additive steps	No mathematics Reason	
Following Monday	Delayed checks on the additive steps	Lesson 8: related facts to 10 × 10, and the review of learning	

Note. Days and sessions follow the example ARC week in Table 4. Lesson titles are paraphrased from the Department's unit, and the Department's Lesson 1 on grid maps is moved to Thursday so that it can be taught outside.

Session plan 3: Augment, Stage 2 mathematics, Week 2 Wednesday, 12:00 to 12:40

The step is choosing an efficient strategy for addition and subtraction (MA2-AR-01, MAO-WM-01). About 26 students work with two teachers.

Time	Move	What a student working at Stage level meets
12:00	Retrieve	Six mixed items from secured steps: two number facts, a change problem, and a delayed check on last week's place value step
12:05	Teach	A worked example of levelling: 28 + 35 becomes 30 + 33, shown on the same number line the teacher used on Monday
12:09	Try together	Compensation and constant difference with the hint ladder, as in the exchange in Section 11.6
12:15	Stand and stretch	
12:16	Practise	Mixed problems at about four in five. First the strategy is named. Then the tutor shows a problem and asks which strategy fits before any calculating. Prices come from the class's own seedling price list
12:28	Explain	"For 199 + 56 I would use … because …", then one follow-up: would that work as well for 143 + 56, and why not?
12:32	Check, or carry on	Students who are ready start the ten-problem check, which can finish on Thursday
12:38	Close	One line in the learning log

What the tutor changes for other students. A student who is not yet secure in bridging to ten gets a shorter unit lane, with compensation on two-digit numbers and a number line on screen, and spends the edge lane on number bonds and bridging. A student already secure in all three strategies works with four-digit numbers, meets problems where two strategies tie and has to argue for one, and writes a problem for a classmate in which constant difference is clearly the best choice.

What the teachers do. Teacher A takes a same-wall group of four who compensate in the wrong direction, turning 126 − 58 into 126 − 60 and then taking away two more. They work it through with a number line and counters at the table for ten minutes. Teacher B samples last week's place value checks in three-minute conferences: show me 3,400 in hundreds, and tell me why.

Session plan 4: Reason, Stage 2 mathematics, Week 2 Monday, 12:00 to 12:40

This is the Department's Lesson 4, taught to a class of about twelve with the screens off. It fits in 40 minutes because the independent practice has moved to Augment.

1. **Beginning, 8 minutes.** Review the strategies poster. Students name the strategy shown in three worked boxes on mini whiteboards and give their reasons.
2. **Two problems, 7 minutes.** Pairs solve 27 + 34 and 126 − 58 any way they can, then the class compares methods and names them.
3. **Closest to 100, 20 minutes.** Pairs turn over six playing cards, make two- and three-digit numbers and add or subtract to land as close to 100 as they can. In the first two rounds only the three strategies are allowed, and students record their thinking.
4. **Concluding, 5 minutes.** Which strategy was most efficient, and when is it not?

What travels between the sessions. From Reason to Augment: the models and the words, so the tutor's number line looks like the teacher's. From Augment to Reason: Wednesday's error patterns, which decide what Thursday's lesson opens with. From both to Create: at the fundraiser, students give change at the seedling table in their heads.

Evidence for the unit. The mastery record shows each step as in progress or secured. The teachers' conference notes record the sampled checks. The work samples are photographs of the whiteboards from Closest to 100, which is the Department's own optional assessment for the lesson, and the grid maps.

A.3 The pilot coaching course

These two units are the proposed first build. A pilot coaching course for them would contain five things.

- A syllabus map for each unit. About 25 to 30 steps for the mathematics unit and about eight for the skill-based content of the English unit, each tied to its outcome and content point.
- An item bank written and checked by teachers. For each step: one worked example, a hint ladder, practice items at three levels of difficulty and two forms of the check.
- Tutor scripts for the six moves, including the sentence starters for the Explain move.
- The teacher view with its three lists, and a plain-language family view.
- The paper resources for the Reason sessions and the hands-on stations.

The pilot itself would be one Stage 1 class and one Stage 2 class for a fortnight each, with consent. It would measure the steps secured and still held three weeks later, minutes at the screen, flags per session, teacher minutes in conferences, and what students and teachers say about it. The comparison would be the Department's optional assessments for the same lessons, taught conventionally in a parallel class. That is small enough to build with existing tools, and specific enough to fail.

Licence

© 2026 Dane Smale. Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0). You are free to share and adapt this work for any purpose, including commercial use, provided you credit the author and link to the licence: https://creativecommons.org/licenses/by/4.0/

Suggested citation: Smale, D. (2026). *The PrimarAI School: A manifesto and design proposal for an AI-native Australian primary school* (November 2026 ed.). PrimarAI. https://primarai.com.au/