DECO 6500 A3.2

Consent, on Your Terms

A Design Fiction about Defaults-as-Governance

Pathway 2 · Human–Technology System Futures

Explanatory Memorandum

Rongyang Jian

S4950018

Framing & Intent

This design fiction asks a pointed "what if": what might happen in school mental-health triage if **consent were treated as a process** (revocable, time-bounded, legible) and **defaults were made visible** so that governance by interface could be inspected rather than hidden? The artefact is deliberately **in-world**: a working single-page product UI that behaves as if it already exists, so readers can reason with concrete flows (consent, results, referral/withdrawal) rather than abstractions. The intent is **provocation, not prediction**—to surface how small interface choices (default routes, expiry timers, equal-weight withdrawal) redistribute agency, trust, and workload across students, caregivers, and institutions. That stance follows design-fiction practice as articulated by Dunne & Raby: speculative artefacts are **not forecasts** but **instruments to spark debate** about near-future social-technical relations (Dunne & Raby, 2007). It also adopts Bleecker's **diegetic prototype** move: placing a thing inside a plausible world so audiences can discuss what it **does** to that world (and what the world does back to it), rather than discussing ideas in the abstract (Bleecker, 2022).
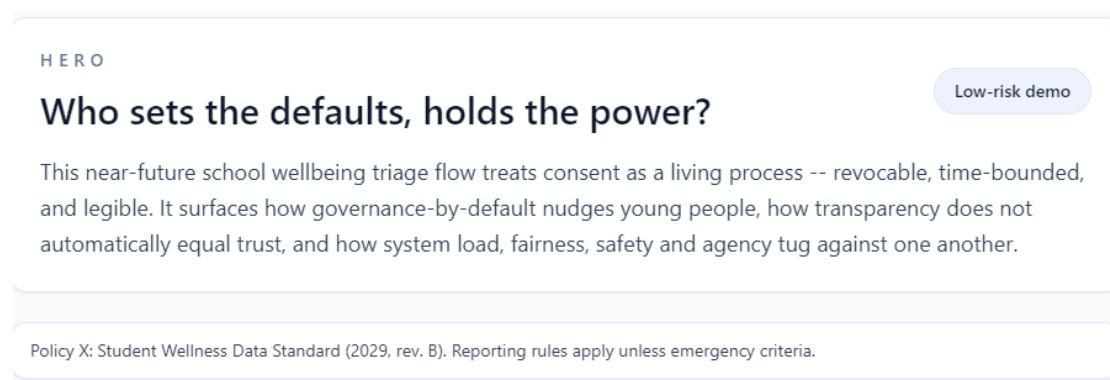

Figure 1 In-world stance; speculation, not prediction.

From A2 to A3: Evidence-to-World Shift

A3 extends four A2 observations into system rules that people can experience in-world.

- **Hesitation around "who will be told."** In A2, participants paused at disclosure moments. A3 turns that friction into editable defaults: **chips** expose "Default: Notify later" and "Consent expires: 30 days," with immediate **Edit**. This makes **choice architecture** auditably visible, rather than nudging invisibly.

Figure 2 default chips + unchecked "screening ≠ diagnosis"

- **Usability acceptable, trust fragile.** A2 scores (e.g., SUS ≈ 67; UTAUT mid-high) suggested functionally usable flows that still felt risky. A3 responds with **neutral risk bands**, a **Why-this-score** panel that enumerates **three evidence types** (item clusters; self-report; time-trend), and a printable **consent receipt**—practices aligned with recognized explainability guidance for decisions that affect individuals.

Figure 3 Why-panel open and a downloaded consent-receipt

- **"Red risk" visuals increased anxiety.** A3 replaces alarm colors with **ranges** and **last-updated timers** for queue estimates, acknowledging uncertainty instead of performing certainty.

QUEUE STATUS
moderate · wait 15-25 min
Last updated 22:33s

Estimates are ranges, not guarantees.

Route: moderate

Talk now	Book later	Ask anonymously
high demand · 25-40 min	moderate · 15-25 min	separate · 5-10 min

Figure 4 queue during skeleton + "Last updated" timer

- **One-way reporting suppressed help-seeking.** A3 makes **withdrawal as easy as giving consent** (equal-weight buttons + 2-second feedback that alerts are cleared and queue slots recycled). This operationalizes the principle that withdrawal should be as easy as consent.

SCREEN 3

Consent is revocable

Referral / Withdraw consent
(1) Who → (2) When → (3) Confirm

Who is notified?

☑ Counsellor

☐ Parent / Guardian

☐ Year level lead

When?

○ Now

◉ End of day

○ Choose time

↩ View / Withdraw consent   Confirm →

Policy note: Selecting Parent follows school policy X; your choice remains respected unless emergency criteria. Emergency criteria override consent per Policy X §4 (harm/immediate risk).

Figure 5 equal-weight Withdraw/Confirm and the 2-second feedback toast

HCI Critique & Theoretical Grounding

The fiction examines **defaults-as-governance** and **consent-as-process**. Choice architecture research shows that defaults quietly steer behavior; treating them as "neutral" hides power. Our UI therefore surfaces and lets learners edit defaults (e.g., notification route, expiry), making the nudge legible and accountable. We pair this with **revocable, time-bounded consent** operationalized as equal-weight "Confirm/Withdraw" and a machine-readable receipt. This follows the legal principle

that **withdrawal must be as easy as giving consent**, and aligns with **Privacy by Design** where protective settings are the default, not an advanced option. Together, these moves challenge the industry habit of embedding governance in fine print or invisible toggles (Chapman et al., 2021).

The fiction also contests the **transparency fallacy**—the belief that showing numbers equals trust. Ananny & Crawford caution that transparency alone often fails as accountability; people may "see" yet not genuinely "know" (Ananny & Crawford, 2018). We therefore favour **seamful** elements that reveal system limits rather than simulate certainty: neutral bands (not alarm colours), **ranges** for queue times, last-**updated** timers, and a **Why-this-score** panel that lists evidence types learners can understand. This implements practical explainability guidance (ICO × Alan Turing Institute) that centres affected individuals and context, not just model internals (ICO., 2025). The result is a critique turned into components: defaults you can inspect, consent you can revoke, explanations you can follow, and seams that teach uncertainty.

System Map: Roles, Rules, and Feedback Loops

Actors. Student; counsellor; parent/guardian; year lead; school admin; policy layer (e.g., Student Wellness Data Standard).

Rules. Who gets notified; when (now/later); consent **expiry; withdrawal** parity; emergency override; documentation via consent **receipts**.

State. A triage **queue** with finite capacity; alert registers per stakeholder; a consent ledger with versions.

Loops.

1. **Revocation loop.** Student withdraws → alerts cleared → queue slot recycled → system load shifts. This makes consent a live control, not a one-time gate, and turns governance into an ongoing negotiation visible in UI and logs.
2. **Uncertainty loop.** Traffic changes route estimates: UI shows **ranges + last-updated**; users adapt plans; subsequent demand changes again. The system embraces uncertainty, avoiding false precision that erodes trust.
3. **Default loop.** Exposed defaults (e.g., "notify later", 30-day expiry) affect behavior; revised default policies feed back into institutional outcomes.

Leverage. As Meadows notes, small changes in **information flows and rules** are high-leverage interventions in complex systems: make defaults visible/editable, make

withdrawal operational, and make uncertainty explicit (Donella, n.d.). Our map pinpoints these levers at the interface and policy boundary so that shifting them predictably changes stakeholder experience and workload.

Two Trajectories: Undesirable ↔ Desirable

Undesirable. Consent is treated as a one-off checkbox; withdrawing is obscure or burdensome, so alerts persist and data keeps flowing. "Defaults" silently govern behavior, locking students into notify-now pathways that maximize KPIs rather than care. Interfaces perform *transparency theatre*—red risk labels and single-number ETAs that look authoritative but conceal uncertainty and operational limits. In emergencies, overrides creep from rare exceptions to routine practice, normalizing surveillance while eroding agency. These patterns are predictable: choice architecture makes defaults powerful, and "just show the numbers" rarely produces accountability or trust.

Desirable. Consent is a **process**: equal weight Confirm/Withdraw, printable receipts, editable expiry windows. This makes withdrawal as easy as giving consent—not only in policy text. Defaults are **visible and editable** (chips), so governance by interface becomes inspectable. Transparency is replaced with explainability-with-seams: a "Why this score?" panel listing evidence types; queue ranges with last-updated timers; neutral bands instead of alarm colors. Communicating uncertainty as numeric ranges preserves credibility better than pretending certainty, and **seamful** elements teach system limits rather than hiding them (Chalmers, 2003). Privacy-by-Design principles tie it together: protection as the default, not an advanced option.

Design-Decision ↔ Critique Mapping

- **Defaults-as-governance → Default chips + editable expiry.** Surfaces the nudge so it can be inspected or changed (choice architecture made visible).
- **Consent-as-process → Equal-weight Confirm/Withdraw** + machine-readable receipt.
- **Transparency fallacy → "Why this score?" lists evidence types.** Shifts from raw numbers to accountable, person-centered explanations.
- **Uncertainty communication → Queue *ranges* + "last updated" timers.** Avoids false precision; numeric ranges maintain credibility better than vague disclaimers.
- **Seamful design → Expiry prompts, skeleton-loading, route changes that**

reveal load. Makes infrastructural limits legible and learnable.

- **Privacy by Design → Default-to-privacy notification routes + receipts.**
  Positions restraint as default, escalation as explicit choice.

Scope, Risks, and Mitigations

Scope. This is an in-world speculative artefact to interrogate near-future HCI choices; it is not a clinical device or forecast. Data are simulated; no personal data are processed.

Risks. (i) Misreading ranges as guarantees; (ii) Emergency override being normalized; (iii) Consent parity ignored in deployment.

Mitigations. (i) Numeric ranges + "last updated" timers and copy that uncertainty is expected; (ii) explicit policy note that overrides are constrained; (iii) withdrawal parity implemented in UI and receipts, aligned with GDPR Art. 7(3), and Privacy-by-Design defaults(Van Der Bles et al., 2019).

Art. 7 GDPR
Conditions for consent

1. Where processing is based on consent, the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data.

2. [1] If the data subject's consent is given in the context of a written declaration which also concerns other matters, the request for consent shall be presented in a manner which is clearly distinguishable from the other matters, in an intelligible and easily accessible form, using clear and plain language. [2] Any part of such a declaration which constitutes an infringement of this Regulation shall not be binding.

3. [1] The data subject shall have the right to withdraw his or her consent at any time. [2] The withdrawal of consent shall not affect the lawfulness of processing based on consent before its withdrawal. [3] Prior to giving consent, the data subject shall be informed thereof. [4] It shall be as easy to withdraw as to give consent.

4. When assessing whether consent is freely given, utmost account shall be taken of whether, *inter alia*, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary for the performance of that contract.

Figure 6 GDPR Art. 7(3)

source: https://gdpr-info.eu/art-7-gdpr/

Evaluation Signals

This fiction "works" if audiences can, unaided, restate the critique as **defaults configure power; consent is an ongoing negotiation; transparency without seams doesn't yield trust.** Look for: (a) readers naming defaults and expiry as governance; (b) recognition that **withdrawal parity** changes behavior; (c) acceptance of ranges and last-updated as honest uncertainty (no perceived drop in credibility); (d) calls to adopt receipts/Why-panels in real services (Van Der Bles et al., 2019). Empirically, uncertainty expressed as numeric ranges tends not to reduce trust compared to overconfident point values—so feedback endorsing the ranges is a positive signal.

References

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, *20*(3), 973–989. https://doi.org/10.1177/1461444816676645

Bleecker, J. (2022). Design Fiction: A Short Essay on Design, Science, Fact, and Fiction. In S. Carta (Ed.), *Machine Learning and the City* (1st edn, pp. 561–578). Wiley. https://doi.org/10.1002/9781119815075.ch47

Chalmers, M., & MacColl, I. (2003). *Seamful Design and Ubicomp Infrastructure*[Workshop paper]. Workshop at the Crossroads: The Interaction of HCI and Systems Issues in UbiComp, Ubicomp 2003. https://www.dcs.gla.ac.uk/~matthew/papers/ubicomp2003HCISystems.pdf

Chapman, G., Milkman, K. L., Rand, D., Rogers, T., & Thaler, R. H. (2021). Nudges and choice architecture in organizations: New frontiers. *Organizational Behavior and Human Decision Processes*, *163*, 1–3. https://doi.org/10.1016/j.obhdp.2020.04.004

Donella, M. (n.d.). Leverage Points: Places to Intervene in a System. *The Academy for Systems Change*. Retrieved 8 November 2025, from https://donellameadows.org/archives/leverage-points-places-to-intervene-in-a-system/

Dunne, A., & Raby, F. (2007).https://dunneandraby.co.uk/content/bydandr/36/0

Information Commissioner's Office & The Alan Turing Institute. (2020). *Explaining*

decisions made with AI. ICO. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/

Van Der Bles, A. M., Van Der Linden, S., Freeman, A. L. J., Mitchell, J., Galvao, A. B., Zaval, L., & Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. *Royal Society Open Science*, *6*(5), 181870. https://doi.org/10.1098/rsos.181870

Appendix A. A2 Quantitative & Qualitative Summaries
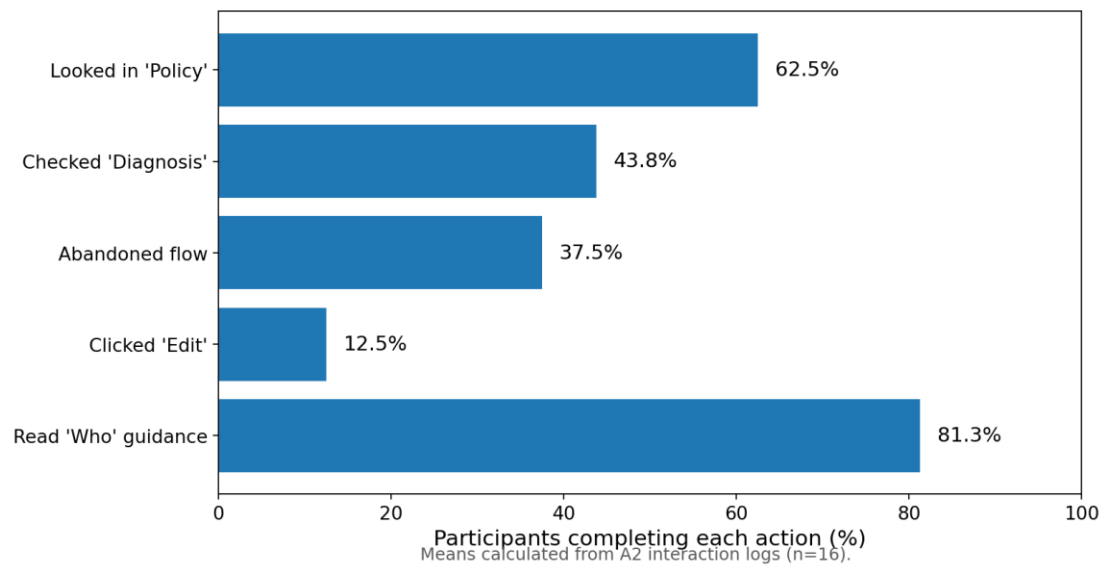
A2 Evidence: Pause at Disclosure

Figure A.1 Users Pause at Disclosure

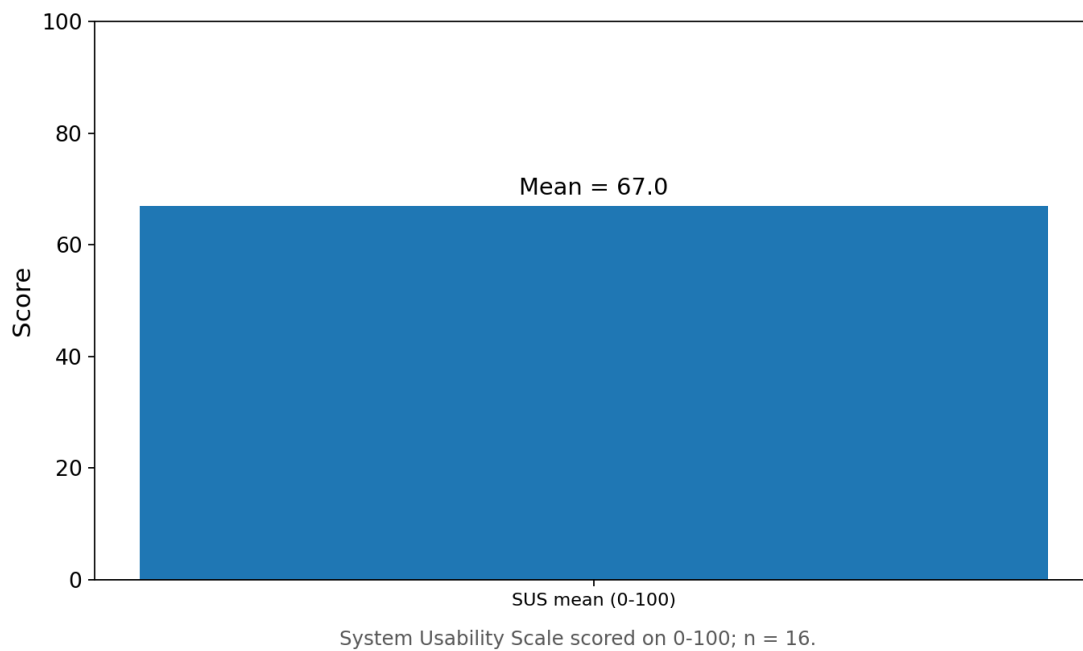## A2 Evidence: SUS Summary

Figure A.2 System Usability Scale (SUS) summary
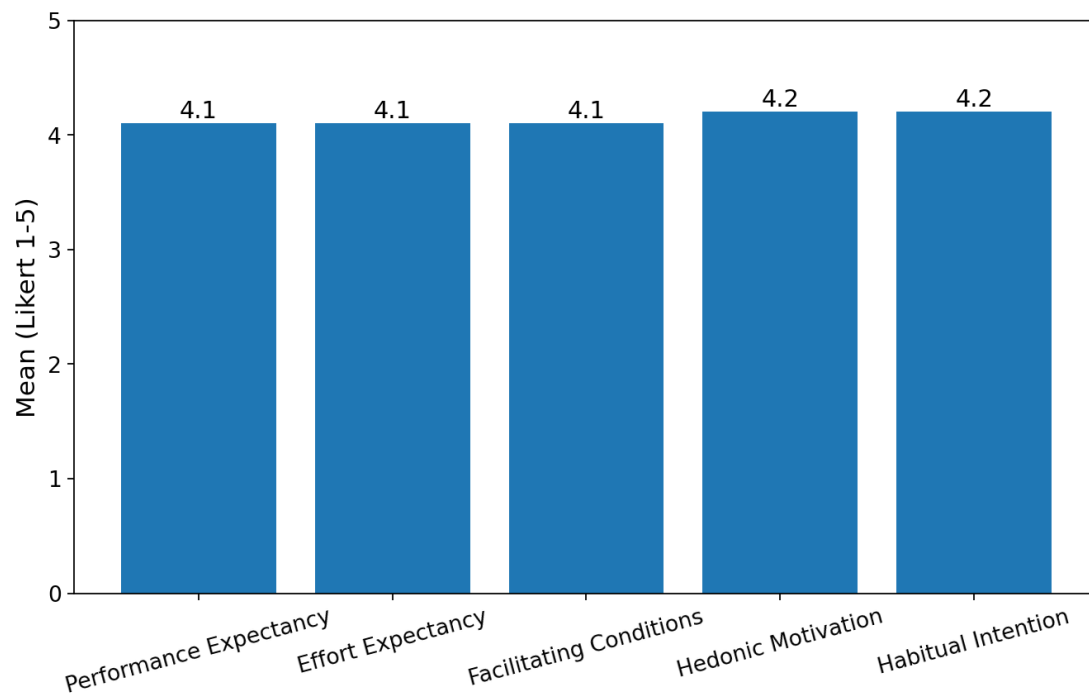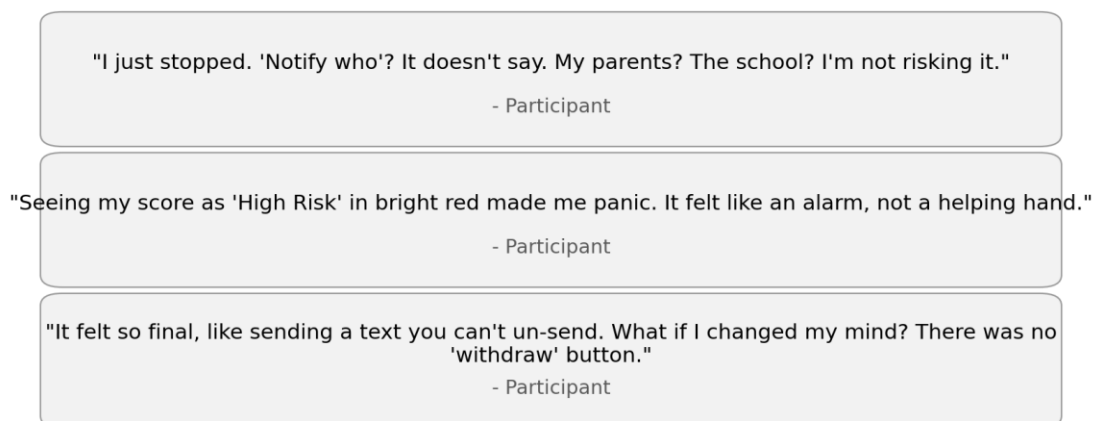
A2 Evidence: UTAUT Subscales

Figure A.3 UTAUT subscales

A2 Evidence: Participant Comments on 'Red risk' Visuals

"I just stopped. 'Notify who'? It doesn't say. My parents? The school? I'm not risking it."

- Participant

"Seeing my score as 'High Risk' in bright red made me panic. It felt like an alarm, not a helping hand."

- Participant

"It felt so final, like sending a text you can't un-send. What if I changed my mind? There was no 'withdraw' button."

- Participant

Participant reactions highlight anxiety triggered by red 'High Risk' cues and desire for reversible consent actions.

Figure A.4 Comments on "red risk" visuals

Appendix B. Connect to A2 insights

Systemic Insights (1)

(A1 reflection)

A2 Insight

Students hesitate or drop off at the consent step; how consent is framed and timed changes whether they reach a human

Evidence (quant and qual data including numbers and quotes)

SUS ≈ 67.3/100 (marginal/acceptable); UTAUT-2 means - PE 4.1, EE 4.1, FC 4.25, HM 4.2, HT 4.19

Observed pattern (usability + think-aloud): abandonment clusters around consent (step 2); students scan the screen to see "who will be told" before proceeding.

Why this insight matters

Consent is the ~~virtued~~ valve that governs early disclosure. Lowering privacy ambiguity without removing safeguards increases student-initiated contact and trust.

Level of this insight:
- Interface
- Systemic
- Both
Why?

Both.
Interface copy, defaults, and micro-interactions set the friction, but the rule set (who is notified, when, and by whom) is institutional.

Futures Grid

Insight
What pattern, feedback loop, or stakeholder tension emerged in A2?

Why privacy is ambiguous, student abandon at the consent step when consent is granular, time-locked, and clearly explained, contact increases Dynamic: a balancing loop "privacy friction ↔ outreach" with leverage at rules and information flows. stu, parent, counselor, teacher tensions)

Desirable Future/s
What if this insight were amplified in positive ways? Who benefits?

1. Granular, time locked consent with downloadable" consent receipt, leads to earlier partial disclosures and faster counselor contact; students and counselor benefit from reduced drop-offs. 2. Microcopy '微文案' + a one-tap comprehensio check '"screening ≠ diagnosis", raises understanding and trust; students and parents benefit from clearer expectations. 3. load-aware intake shows safe routing (now/later/anon) at the consent step; queues smooth out and Schools benefit from steadier workload.

Undesirable Future/s
What if this insight produced negative or unintended consequences? Who is harmed or excluded?

1. too many toggles creates choic overload; at-risk stu freeze and drop-off
2. urgent cases are hidden from guardians too long; increase harm and liability
3. Complex language disadvantages (non-English speaker); equity gap widens as completion rates fall for these groups.

Critical HCI Tensions
What principles, theories, or values are challenged or need to change?

1. Safety vs. Agency
2. Simplicity vs. Granularity: fewer choic vs. meanful revocable choices
3. comprehension vs. burden. consent clarity vs. drop off

Pathway

(3)

To complete this, you will need to review the information in the Task Sheet.

The pathways are outlined on Pages 3 and 4.

My selected pathway is ...

Pathway 2:

Human - Tech

My rationale for selecting this pathway is ...

user's direct experence with interface text and consent is where this core conflict is felt.

2 is allow us to demonstrate that "transparency ≠ trust" and to explore the critical role that withdrawable consent and privacy-by default play in the difficult decision to seef - help.

Framing Statement ④

My potential design fiction will focus on ...

A Student moment of consent and result-interpretation in a school mental-health help-seeking apps, showing how microcopy default settings, revocable, time-locked consent shape whether they reach a human counselor

It matters because ...

getting help early really depends on the consent process. If that process is confusing or unclear, people lose trust and will longer to get help.

But if choices are simple, the rules are obvious, and people know they can easily change their minds later, drop-off rates decreases

The key HCI critique is ...

Showing peoples info, like risk score or a legal warning, Does not build trust or give them real control.

the system must design from: 1. Be private by default, don't have to search setting

2. Let user easily change their mind (revocability)

3. easy be understanding, not confusing or overwhelming

The key systemic issue is ...

When the school's target is hitting number, "find students who need help early, that goal starts to control the app's design. means school's target - not stu - ends up deciding the setting for who agree what, and who get notified

When this happen, stu lose control over their own personal info make this system unsafe and unfair for everyone involved.

Appendix C. Activity 1: Pathway Alignment Worksheet

Purpose:

To help you decide which of the two A3 pathways best aligns with your topic, and to clarify what kind of speculative artefact might communicate your critique.

Your staring point

Write one or two sentences summarising the systemic insight or tension you identified from your A2 wo

On campus mental health screening apps, the combination of confusing disclaimers, unclear consent forms, and pressure to meet performance goals makes students feel unsafe and less likely to trust the platform. This stops them from seeking help early. When rolled out on a large scale, this leads to greater privacy risks and overwhelms support services, as safety rules are ignored in Favor of hitting targets.

Compare the two pathways

	Systemic Evaluation Futures	Human–Technology System Futures
Scope / Focus	Systems, infrastructures, policies, institutions	Human experiences, interactions, identities
Heuristic 1 – Where does your critique live?	☐ Is the problem in how systems are structured, governed, or measured?	☑ Is the problem revealed through how people experience or interpret technology?
Heuristic 2 – What changes in this world?	☐ What if the rules, policies, or metrics of the system changed—how would society adapt?	☑ What if the interfaces or relationships changed—how would people feel or behave?

Heuristic 3 – What must your artefact *do*?	☐ Expose or critique institutional logics (e.g. report, dashboard, policy memo).	☐ (highlighted) Reveal lived experiences (e.g. diary, ad, UI mock-up, testimonial).
Heuristic 4 – Scale of Consequence	☐ Ripples across multiple organisations or communities.	☐ (highlighted) Plays out in daily life or personal identity.
Heuristic 5 – Type of Critique	☐ Values, control, power, governance.	☐ (highlighted) Relationships, perception, agency, ethics of interaction.
Possible Formats	Annual report • policy memo • news article • service blueprint	Ad • interface demo • first-person narrative • user diary
Notes		

Quick Diagnostic

Tick all that apply

☐ √ Needs a **character** or lived experience → *Human–Technology (pathway 2)*

☐ Needs a **policy** or system → *Systemic Evaluation (pathway 1)*

☐ √ Asks "What would it feel like?" → *Human–Technology (pathway 2)*

☐ Asks "What would it mean or cause?" → *Systemic Evaluation (pathway 1)*

Decision and reflection

I will likely pursue the [Systemic Evaluation / Human-Technology] pathway because …

My choice: Pathway 2 — Human–Technology

The user's direct experience with the interface text and consent process is where this core conflict is felt. That's why Pathway 2 is essential—it allows us to

demonstrate that "transparency ≠ trust" and to explore the critical role that

withdrawable consent and privacy-by-default play in the difficult decision to seek help.

Possible artefact format

A set of annotated design mockups for the app's interface. Specifically, it would show:

- The Screening Results Page: What students see after they finish the questionnaire.

- The "Why this score?" Pop-up: The platform's attempt to explain the results.

- The User Path for Critical Actions: The step-by-step process for the "one-click referral," how it notifies parents, and the option to withdraw consent.

HCI principle/value my fiction will challenge

Transparency vs. Trust.

Agency & Revocability.

Privacy-by-default.

Explainability/Interpretability.

Metrics-as-meaning

Appendix D. Activity 2: World-building Worksheet

Select a key HCI tension from your futures grid	On campus mental health screening platforms, the communication of the critical boundary where "screening is not diagnosis"—along with the default settings for consent and notification—is driven by KPIs like the "early identification rate." This leads to a situation where transparency fails to build trust, ultimately altering when students seek help and the privacy risks they face.

Build your world using the five lenses

Actors	Institutions & Structures	Norms & Values	Technologies & Interfaces	Faint Signals
Who are the key players? Who gains or loses? Whose perspective matters	What rules, organisations, or infrastructures define this world? Who enforces them?	What behaviours or beliefs are normal? What is taboo? How do people justify the system?	What technologies are "everyday"? How do people interact with them? What affordances shape life?	What small trends, prototypes, or news items suggest this future is emerging?
Students (14–18) Parents/Guardians Homeroom Teachers / Heads of Year School Counselors / Psychologists	State Department of Education evaluations and funding allocation. School governance structures and crisis intervention protocols. Privacy and data retention policies. Vendor compliance and Service Level	The consensus that "proactive screening is superior to passive waiting." The assumption of guardianship where data is visible to parents by	Core Tech: School SSO integration, questionnaire scales +	Growing public discourse on "consent fatigue" and the importance of revocable consent.

School Data Officer / Dean of Students (or Head of Moral Education) Platform Product Managers & Algorithm Team External Mental Health Services / Hotlines State Department of Education & Funding Evaluators Third-party Auditors / Appeals Channels Beneficiaries: Schools that meet their targets, platform vendors. Those Harmed: Students who are falsely flagged or whose privacy is exposed by default.	Agreements (SLAs). Evaluation Metrics: Early identification rate, referral completion rate, response time. Implementers/Enforcers: School administration, data officers, third-party auditors.	default; the dissemination of Red/Yellow/Green risk classifications. The intuitive belief that "more information = more transparency = more trust." Taboos: Publicly identifying students who withdraw consent or refuse to be screened; teachers questioning the KPIs. Rationalizations: Justifying the risks associated with default settings and outsourcing in the name of "safety, efficiency, and compliance."	lightweight models, cloud-based calibration, and activity logs. User Interface: A "Why this score?" explainability component on the results page; one-click referral and scheduling interfaces. Control Mechanisms: A parent notification toggle (on by default); an entry point to "withdraw and delete data" (deeply nested and visually de-	An emerging micro-trend of "digital minimalism" or "refusal to be screened" among students; news of campus data breaches is triggering a crisis of trust. Continued strain on the capacity of mental health services; "explainability labels" are progressively being incorporated into regulations.

			emphasize d).	

World Title	Metrics of Care
Give your world a headline or nickname that captures its mood of theme	