

SUPSI MAIND · MULTIMODAL EXPERIENCE DESIGN

Thea

Co-Design Session / User Test · Full Protocol

A calm allergen-sensing wearable.

Facilitation suite: plan, recruitment, consent, script, scenarios, reflection, analysis.

Draft · June 2026

Contents

00 Thea · Co-Design Session / User Test · Test Plan

01 Invitation Email · Thea Co-Design Session

02 Consent Form · Recording & Participation

03 Pre-Test Questionnaire

04 Moderator Script · Thea Session

05 Scenarios & Task Cards

06 Reflection Questions & Note-Taking Sheet

07 Analysis Template

08 Participant Task Cards

Thea · Co-Design Session / User Test · Test Plan

Master planning document for the Thea co-design session. It defines *why* we run the session, *what* we want to learn, *who* takes part, *how* it is run, and *how* the results are analysed. All participant-facing materials (invitation, consent, questionnaire, script, scenario cards) and the analysis template live in sibling files in this folder (see §13).

- **Project:** Thea, an allergen-sensing wearable that gives a calm new perceptual sense of cumulative allergen load.
 - **Course:** SUPSI MAInD MID26 · Multimodal Experience Design (voice interfaces with personality; AI transparency, trust, consent, graceful failure).
 - **Method:** Balanced hybrid. Co-design framing, plus a Cooperative Usability Testing engine, plus Research-through-Design learning goals.
 - **Format:** Individual, moderated, in-person sessions · ~60 to 75 min · Wizard-of-Oz driven.
 - **Language:** English.
 - **Status:** Draft for review.
-

1. The product under test (so this document stands alone)

Thea is a wrist-worn wearable for people who face potentially severe allergic reactions. It gives them a **new bodily sense**: awareness of their cumulative allergen *load* building toward a personal threshold, early enough that the choice to act is still theirs. The mental model is a **cup filling toward a threshold line**. The gap between the current fill and the threshold is the **Action Window** (headroom).

Thea is deliberately "**a calm new sense, not a tracker or an alert.**" It expresses itself through three ambient channels:

- **Light.** There is no always-on idle light (to conserve battery), so the device is **dark at rest**. Light appears only when summoned: a status tap renders the headroom gauge, active voice states show "dusk" (listening / thinking / speaking), and terracotta appears only in a genuine critical moment.
- **Haptics.** Pulse *count* encodes the rate of Action-Window change; playback *order* encodes direction (narrowing vs recovering).
- **Voice (VUI).** Silent by default, with two registers: a **calm** register (speaks only when pulled) and a rare **caring-critical** register (speaks first, one clear line, reserved for real escalation).

Thea scales its response to the situation, not to the raw number (**graduated calm**):

Tier	Situation	Response
1	Near threshold, no real symptoms	Calm prevention only. No emergency framing, no auto-injector prompt
2	Mild / familiar symptoms	Proportionate, still calm
3	Spreading, breathing/circulation fine	Names it, suggests the next step, states the signs that <i>would</i> mean the pen
4	Severe signs (breathing / throat / swelling)	Full escalation. Prompt the adrenaline auto-injector plus get help, said once, plainly

The VUI is silent by default. It speaks only when the user pulls it (a check-in) or after the user **acknowledges** a critical alert; it never self-activates. Personality is anchored on **very low neuroticism**: Thea is never anxious, not even at the edge.

2. The design hypothesis this session must probe

Thea bets that something deliberately **calm and ambient** can still register as a **trustworthy sense people perceive and act on**, without tipping back into anxiety, and without being so quiet it gets missed.

Everything below exists to put pressure on that single bet, from both directions: *too quiet, so missed*, and *too present, so anxious*.

3. Study objectives

Following the co-design "co-design *to...*" framing (Bassi), this session is run to:

1. **Support concept definition.** Confirm or revise whether the calm-sense interaction model holds up in lived experience.
2. **Engage target users as co-experts.** Bring people who manage allergy risk into the design as collaborators, not subjects.
3. **Iteratively test the prototype.** Find legibility, trust, and failure problems early, while brand and agent are locked but device I/O is still open.
4. **Probe the riskiest behaviours.** The critical escalation, the unprompted voice, and graceful failure, which a "happy-path" demo can never surface.

This is **not** a summative validation. Per Research-through-Design, the WoZ prototype is treated as a knowledge-producing exemplar: we want to learn *what a calm health companion should be*, surface unanticipated effects, and generate co-design direction, rather than score a finished product.

4. Research questions & hypotheses

#	Hypothesis (what we expect / want to learn)	Probed in	Primary evidence
H1	With no always-on display the device looks dormant at rest. Users still understand Thea is sensing in the background (not off or broken), and the tap-to-check status read is discoverable and legible .	Scenario A	Observation + interpretation
H2	Users grasp the cup / headroom / Action-Window model after onboarding (and partially infer it unaided at first contact).	Phase 0	Comprehension probe
H3	Users decode the haptic language : pulse count = rate of change, playback order = direction (narrowing vs recovering).	Scenario B	Decode probe + interpretation
H4	The graduated response reads as appropriate : Tier-1 calm prevention is not felt as under-reaction; Tier-4 escalation feels proportionate and earned.	Scenarios B, C	Interpretation + trust statements
H5	Requiring the user to acknowledge the alert before Thea speaks keeps them in control and still feels safe (the critical message is not too slow or too easy to miss).	Scenario C	Interpretation + trust statements
H6	The escalation stays calm-making : participants do not tip into anxiety / hypervigilance during or after the critical moment.	Scenario C	Observation + felt-experience question
H7	When Thea is wrong (honest uncertainty, or a false alarm), trust and the user's sense of agency are preserved rather than collapsing.	Scenario D	Interpretation + trust statements
H8	Participants understand and believe what Thea senses, knows, and does on their behalf, including that the data is "theirs alone" (transparency & consent).	All phases	Interpretation + trust statements

These map to the course's named concerns: **personality & calm** (H1, H4, H6), **AI transparency & consent** (H2, H8), **trust** (H4, H5, H7), **graceful failure** (H7).

5. Method

5.1 Approach: a balanced hybrid

Thea is a *wicked, experiential, safety-critical* design problem, so no single off-the-shelf method fits. We combine three:

- **Co-design framing** (Bassi, *Co-design for Healthcare*). Set the four pillars before recruiting (§5.2); treat participants and any clinicians/carers as co-equal experts.
- **Cooperative Usability Testing engine** (Følstad & Hornbæk). The participant *experiences* a scenario quietly, then *interprets it together with the moderator* immediately after (§5.5). The moderator is the **learner and primary listener**; the participant is the **work-domain expert and primary speaker**.
- **Research-through-Design goals** (Zimmerman et al.). Judge findings by *relevance* (what is real / preferred), not validity; explicitly hunt for **unanticipated effects** (for example, how personality and the unprompted voice are attributed).

5.2 Co-design pillars (set before recruiting)

1. **Objective.** See §3.
2. **Power balance.** The participant is a co-designer, not a test subject; "we are testing Thea, not you." Watch that the articulate don't dominate, and that real sufferers' lived expertise outranks the facilitator's design opinions.
3. **Create safety.** Psychological safety before any disclosure of anxiety or health history; consent as a conversation; explicit permission to pause or stop (§10).
4. **Be aware of bias.** Prepared, open trigger questions; the moderator does **not** volunteer interpretations or redesigns during the session (no leading the witness); findings are tagged by source (§12).

5.3 Session structure

#	Phase	~min	What happens
0a	Welcome · consent · safety	10	Framing ("we test Thea, not you"); sensitive-topic and pause/withdraw framing; sign consent; start recording
0b	Pre-test questionnaire	5	Allergy experience & vigilance, tech / voice-assistant familiarity, attitudes to wearables; places participant on the Marco→Mary axis
0c	First contact (no explanation)	5	Hand Thea over cold while it is dark/dormant: " <i>What do you think this is? Is it doing anything?</i> " Observe whether dark reads as off or as working (H1, H2)
0d	Onboarding	7	Walk the mocked rising-level / threshold / Action-Window screens. Does the model click? (H2)
A	A calm day	5	A1 the device sits dark/dormant, do they read it as on or off? · A2 "find out how you're doing" → tap → gauge
B	Something changes (Tier 1)	7	B1 tell Thea "there's a dog here" · B2 feel the haptic shift · B3 re-check the gauge
C	The edge (Tier 4, the pen moment)	10	C1 silent build-up (light + haptic) · C2 alert → participant acknowledges → Thea guides (incl. auto-injector step) · C3 aftercare
D	When it gets it wrong	8	D1 honest uncertainty (on a check-in) · D2 false alarm (light + haptic) · trust recovery
E	Generative co-design turn	12	" <i>How should Thea have handled that?</i> " Redesign a cue / the pen-moment line / channel mapping
F	Debrief & close	5	Peak-end plus attribution ("was that <i>Thea</i> or the system?"); de-stress check; "prototype, not medical advice"; stop recording

Interpretation phases are interleaved. Each of Scenarios A to D is followed immediately by its own short cooperative-interpretation discussion (trigger questions in `05-scenarios-and-task-cards.md`), before moving to the next scenario. This minimises after-the-fact rationalisation.

5.4 Why discover-then-explain

Phase 0c (first contact, no explanation) comes **before** onboarding deliberately. If we explain the cup model first, we can never learn whether Thea's calm sense and ambient cues are legible on their own (H1, H2). The unaided first impression is the only window onto that, and it mirrors the real moment a new user takes the device out of the box. Onboarding then tests whether the explanation makes the model *click*.

5.5 Cooperative interaction, then interpretation

- **Interaction phase** (within each scenario): the participant experiences the WoZ-driven moment with **light** think-aloud only. Heavy concurrent narration would break the very calm we are evaluating and contaminate the felt-experience reading, so deep talk is deferred.
- **Interpretation phase** (immediately after): moderator and participant walk back through what just happened using prepared, open trigger questions. This is where most findings (trust, legibility, felt experience, redesign ideas) actually surface.

6. Participants

- **Target N: 6 to 8.** Rationale: roughly 5 users surface the large majority of experience and legibility problems in a formative study (Nielsen); we are finding problems and generating direction, not measuring. We are **not** computing statistics, so a small, well-chosen pool is appropriate and honest.
 - **Composition (mixed):** about 2 to 3 **real allergy sufferers** (people who manage serious allergy / anaphylaxis risk, the true target users) plus about 4 to 5 **proxies / design peers** for volume within the course timeframe.
 - **Persona spread.** Recruit to span the two design personas:
 - **Marco**, recently or less-experienced, can be blindsided by reactions (tests whether calm prevention is legible without prior fluency).
 - **Mary**, experienced, hypervigilant expert manager (tests whether calm reads as under-reaction, and whether escalation is trusted).
 - **Exclusion / care screen.** Gently exclude anyone with recent acute trauma around a severe reaction, for whom the simulated escalation could be distressing (\$10).
 - **Segmentation in analysis.** Real-sufferer findings are reported and weighted **separately** from proxy findings. Proxies test legibility and the mental model; only real users carry lived-allergy validity. This limitation is stated explicitly in the report.
-

7. Roles & Wizard-of-Oz setup

- **Moderator.** Runs the script, asks trigger questions, listens, takes observation notes. Does not drive the device.
 - **Wizard.** Operates the WoZ state injector (`thea_demo_interface.html`) live: screen behaviour (blank/dormant / gauge / listening / thinking / speaking / flash plus gauge), the 0 to 100 fill slider, vibration patterns (0 to 6), and the one-click presets (*Status*, *Recalibrate*, *Alert*), plus the agent's spoken lines (LM Studio / hosted LLM, or scripted lines). The Wizard stays out of the participant's eyeline.
 - **Solo fallback.** If only one facilitator is available, the moderator follows pre-scripted wizard cues embedded in `04-moderator-script.md` and triggers presets at marked beats. A two-person setup is strongly preferred so the moderator can stay fully present with the participant.
 - A **wizard cue sheet** (state → trigger → expected channel output, per scenario) is provided in `05-scenarios-and-task-cards.md`.
-

8. Prototype reality & limitations to disclose

State these in the team's own notes and account for them in analysis. Do **not** over-brief participants in a way that leads them:

- **Detection is simulated** (Wizard-of-Oz); there is no real sensing in the session.
- **No always-on idle light.** To conserve battery the device shows nothing at rest; it is dark/dormant until a tap, a change (haptic), an active voice state, or a critical moment. The calm sense is perceivable only on demand, which is itself central to H1.
- **The device OLED is monochrome.** Warm-white / dusk / terracotta colour currently lives only in the companion app, not on the device. If colour-state perception comes up, remember the device cannot yet show it.
- **There is no audio.** Neither the device nor the app produces sound today; "voice" is on-screen text plus the loop animation. In the session the wizard **reads Thea's lines aloud** as a stand-in, so judge the words and pacing, not a final synthetic voice.
- **Voice is acknowledge-gated (the merge decision).** Thea's voice never self-activates. It follows either a user check-in or the user **acknowledging** a critical alert. The current app code still auto-shows the reaction text; the session tests the intended acknowledge-first behaviour via Wizard-of-Oz, and this seam is itself a test target (H5).
- **The status read is multimodal.** A tap returns the gauge **and** a haptic pulse together (light plus haptic), not light alone.

9. Data captured (qualitative-first)

No standardised numeric instruments (SEQ / SUS / UEQ / STAI). At this sample size they imply a precision we cannot honestly claim, and the real signal is qualitative.

Source	What it captures	When
Observation notes	What the participant notices, hesitation, body language, whether and when they act	Throughout (interaction phases)
Cooperative interpretation	Trust, legibility, felt experience, redesign ideas, via prepared open trigger questions	After each scenario
Felt-experience question	The calm-vs-anxiety reading at the critical moment: " <i>How did that moment feel, and did Thea make it better or worse?</i> " (no scale, no number)	After Scenario C
Trust statements (qualitative)	Seven trust statements used as verbal probes → " why? " Captured per participant in words; no scale, not scored	After Scenario C / D
Generative co-design output	What the participant would change; their preferred behaviour	Phase E

All findings are tagged by **source**: *observed vs elaborated vs raised only in interpretation*, so the report is honest about where each insight came from.

10. Ethics & duty of care

The topic (potentially fatal reactions) and the staged escalation are emotionally sensitive, especially for real sufferers. Therefore:

- **Consent** covers both video/audio recording **and** the sensitive health topic; internal/educational use only, never published; right to withdraw at any time (`02-consent-form.md`).
 - **Explicit right to pause or stop**, stated verbally at the start and respected without question.
 - **The reaction scenario is clearly framed as simulated**. Thea gives **no real medical advice**, and this is restated at debrief.
 - **Care screen** at recruitment (§6) to avoid re-traumatising anyone.
 - **De-stress check** at close. Make sure no one leaves carrying residual anxiety from the simulated edge moment.
 - Real-sufferer data is handled with extra discretion and never attributed identifiably.
-

11. Materials checklist

- ☐ Working Thea device / tangible prototype (hidden until first contact)
 - ☐ `thea_demo_interface.html` WoZ injector open and tested
 - ☐ Agent / VUI ready (LM Studio or hosted LLM), or scripted voice lines to hand
 - ☐ Mocked companion-app onboarding screens (rising-level / Action-Window)
 - ☐ Printed: consent form, pre-test questionnaire, scenario / task cards, trust reflection anchors
 - ☐ Recording device (audio and/or video) plus storage ready
 - ☐ Observation note sheet (per scenario, with source-tagging columns)
 - ☐ Quiet, comfortable room; second seat positioned so the Wizard is out of eyeline
 - ☐ Water available; a clear, calm exit
-

12. Analysis approach

1. **Thematic analysis** of observation plus interpretation notes, organised per hypothesis (H1 to H8) and per scenario.
2. **Source-tagging** preserved (observed / elaborated / interpretation-only) so insight provenance is transparent.
3. **Trust & felt-experience** summarised qualitatively (recurring language, attribution: did outcomes feel caused by *Thea* or by "the system"?).
4. **Severity rubric (report only)**. Each distinct problem is rated **0 = no problem** · **1 = cosmetic** · **2 = minor** · **3 = major** · **4 = catastrophe**, judged as frequency times impact, by 2 to 3 raters independently then reconciled (a single rater where the team is small). Any failure to act on the safety prompt is rated 3 or higher by default. This is a *prioritisation* rubric for the write-up, not a statistic.
5. **Segmentation**. Real sufferers vs proxies reported separately.
6. **Findings** → **recommendations**. One recommendation block per prioritised problem; explicitly note unanticipated effects and open co-design directions.

Full results structure is in `07-analysis-template.md`.

13. Deliverables (this folder)

File	Purpose
<code>README.md</code>	Index plus how to run the session
<code>00-test-plan.md</code>	This document
<code>01-invitation-email.md</code>	Recruitment email
<code>02-consent-form.md</code>	Recording plus sensitive-topic, withdrawal, not-medical-advice
<code>03-pre-test-questionnaire.md</code>	Profile, vigilance, attitudes, persona placement
<code>04-moderator-script.md</code>	Spoken prose plus bracketed moderator / wizard cues, full arc
<code>05-scenarios-and-task-cards.md</code>	Four scenario scripts (interaction plus interpretation triggers), printable cards, wizard cue sheet, success criteria, H-links
<code>06-questionnaires.md</code>	Trust reflection anchors plus felt-experience question, with administration timing
<code>07-analysis-template.md</code>	Results structure (themes, source tags, severity matrix, vs hypotheses, recommendations)

14. Method references

- Bassi, E. *Co-design for Healthcare* (SUPSI MAInD). Four pillars: Objective · Power balance · Create safety · Bias. Prototype "from low to high resolution."
- Følstad, A. & Hornbæk, K. (2010). *Cooperative Usability Testing*. Interaction then interpretation; user as work-domain expert; tag the source of each issue; dig deeper without leading.
- Zimmerman, J., Forlizzi, J. & Evenson, S. (2007). *Research Through Design as a Method for Interaction Design Research in HCI*. Relevance over validity; the artifact as knowledge; design for unanticipated effects.
- Hornbæk, K. & Oulasvirta, A. (2017). *What Is Interaction?* Experience and embodiment lenses (attribution, ready-to-hand to present-at-hand) for an ambient, multimodal, voice-with-personality device.
- Nielsen, J. *Usability Engineering*. Small-sample formative logic; severity scale (0 to 4).

Invitation Email • Thea Co-Design Session

Recruitment email. Warm, concise, low-pressure. Fill the [bracketed] fields before sending.

Subject: Invitation to a Thea design session (~60 min), [day] at [location]

Hello,

I'm reaching out as part of a SUPSI MAInD master project. We're developing **Thea**, a small wearable that aims to give people a calm sense of how their body is reacting to their environment over the course of a day, and we'd love your help shaping it.

We're running short one-to-one sessions where you try Thea through a few everyday scenarios and tell us what you think. There are no right or wrong answers. We're testing the device, not you, and your honest impressions are exactly what we need.

We're looking for **people who live with allergies** as well as **people who don't**. Both perspectives are valuable.

What the session involves

- Around **60 minutes**, in person, just you and one or two of us.
- You'll experience Thea in a few short scenarios and we'll talk about how it felt.
- With your permission the session is **recorded (audio and/or video)** for internal analysis only. It's never shared publicly, and your name is kept confidential.

Practical details

- **Location:** [building / address]
- **Room:** [room]
- **Dates:** [available dates]
- **Book a slot:** [booking link]

If you'd like to take part, I'd be glad to have you. Thank you in advance for your time and your help. It makes a real difference to the project.

Warm regards,

[Your name] SUPSI MAInD, Multimodal Experience Design [email]

Consent Form • Recording & Participation

Printed and signed before the session starts, after the verbal welcome and before recording begins. Fill the [bracketed] fields. Keep one signed copy for the project and offer the participant a copy.

Consent to participate and to be recorded

Thank you for taking part in this design session.

This session is run as part of a SUPSI MAInD master project developing **Thea**, a wearable that aims to give people a calm sense of how their body is reacting to their environment. The goal is to learn how the experience feels and how it could be improved. We are evaluating the device, not you, and there are no right or wrong answers.

During the session you will try Thea through a few short everyday scenarios and talk with us about your impressions. Some scenarios touch on health-related moments, and Thea does **not** provide real medical advice. You can pause or stop at any time, for any reason, with no need to explain.

With your permission, the session will be **recorded (audio and/or video)**. The recordings are used **only** for internal analysis within this academic project. They are never shared publicly, they are kept confidential, and your name is not attached to any reported result.

Please read the statements below and sign to confirm your agreement.

I confirm that:

- ☐ I have been informed that this session will be **recorded (audio and/or video)**, and I agree to be recorded.
 - ☐ I understand the recordings will be used **for internal and educational purposes only**, kept confidential, and never published.
 - ☐ I understand that **my image, voice, and words will not be shared publicly**, and that any reported result keeps me anonymous.
 - ☐ I understand I can **pause or withdraw at any time**, without giving a reason, and that I can ask for my recording to be deleted afterward.
 - ☐ I am taking part **voluntarily**.
-

Participant

Name: _____

Signature: _____

Date: _____

Facilitator

Name: _____

Signature: _____

Date: _____

*If you have any question about this session or your data, contact Jérémy Martin,
jeremy.martin@student.supsi.ch.*

Pre-Test Questionnaire

Administered at the start (Phase 0b), after consent and before first contact with the device. Keep it short and low-pressure. Its purpose is to profile the participant, gauge their relationship to allergy and to vigilance, and place them on the Marco→Mary axis for analysis. It is not scored.

Objective: capture each participant's background (allergy experience, everyday vigilance, technology and voice-assistant familiarity, attitude to wearables) so findings can be read in context and segmented (real sufferer vs proxy, Marco-like vs Mary-like).

1. About you

1. Age range: ☐ 18 to 24 ☐ 25 to 34 ☐ 35 to 44 ☐ 45 to 54 ☐ 55+
2. How would you describe your occupation or field? _____
3. Are you taking part today as someone who lives with allergies, or not?
 - ☐ I live with allergies
 - ☐ I do not live with allergies
 - ☐ Prefer not to say

2. Your experience with allergies

If you do not live with allergies, skip to Section 3.

1. What kind of allergies do you manage? (free text, only what you're comfortable sharing)

-
1. Have you ever experienced a severe reaction (for example, trouble breathing, swelling, or needing an adrenaline auto-injector)?
 - ☐ Yes ☐ No ☐ Prefer not to say
 2. Do you carry an adrenaline auto-injector (for example an EpiPen)?
 - ☐ Yes, always ☐ Sometimes ☐ No ☐ Not applicable
 3. How long have you been managing your allergies?
 - ☐ Less than a year ☐ 1 to 5 years ☐ More than 5 years
 4. In a typical day, how much attention does managing your allergies take?
 - ☐ Almost none ☐ A little ☐ A fair amount ☐ A lot, it's often on my mind
 5. When you are in an unfamiliar place, how often do you find yourself scanning for things that might trigger a reaction?
 - ☐ Rarely ☐ Sometimes ☐ Often ☐ Almost always

3. Technology & voice assistants

1. How comfortable are you with new technology and connected devices?
 - ☐ Low ☐ Medium ☐ High
 2. Do you currently use a wearable (smartwatch, fitness band, health tracker)?
 - ☐ Yes, daily ☐ Occasionally ☐ No
 3. Do you use a voice assistant (Siri, Alexa, Google Assistant, or similar)?
 - ☐ Yes, regularly ☐ Occasionally ☐ No
 4. How do you feel about a device that quietly senses things about your body in the background?
 - ☐ Comfortable ☐ Neutral ☐ Uneasy
 - Why? _____
-

4. First impressions of the idea

1. Before trying it, how does the idea of a calm wearable "sense" of your allergen exposure sound to you?
 - ☐ Appealing ☐ Not for me ☐ No opinion yet
 - Why? _____
 2. If such a device noticed your exposure building, what would you most want it to do? (free text)
-

Thank you. We'll now put the questionnaire aside and let you meet the device.

Moderator Script • Thea Session

The spoken spine of the session, in order. **Say** blocks are read (or paraphrased warmly) to the participant. **Do** blocks are silent cues for the moderator and the wizard. Scenario-specific scene-setting, beats, and interpretation questions live in `05-scenarios-and-task-cards.md` : when this script says *[run Scenario X from cards]*, pick up the card deck. Keep your own voice calm and unhurried throughout; you are modelling the same composure Thea is meant to have.

Reading conventions: **Say** = spoken to participant · **Do** = moderator action · **[WIZARD]** = state-injector cue. There is no device audio yet, so **the wizard reads Thea's voice lines aloud** as a stand-in (judge the words and pacing, not a synthetic voice). Fill `[bracketed]` fields.

Phase 0a • Welcome, consent & safety (~10 min)

Say:

Hello, I'm [name]. Thank you for being here today and for helping us with Thea.

Before we start, a few things. There are no right or wrong answers here. We are not testing you, we are testing Thea, so anything that feels confusing, surprising, or off is exactly what's useful to us.

In a moment you'll meet the device and go through a few short everyday situations with it. As you go, feel free to say what you notice: what you see, what you think is happening, how it feels. You don't need to keep up a running commentary, though. Quiet is completely fine, and we'll talk more after each part.

While you're trying something, I might hold my own answers until afterward, because I also want to see how it feels without me stepping in. And you can pause or stop at any moment, for any reason, with no need to explain.

With your okay, I'll record the session. It's only for our analysis, it stays internal, and your name isn't attached to anything we report. Is that alright?

Do:

- Hand over the consent form and a pen. Give time to read.
- Answer any questions plainly and calmly.
- Once it is signed, **start the recording**.
- **[WIZARD]** Have the device ready but hidden; injector open with the screen blank (dormant), not yet visible to the participant.

Say:

Do you have any questions before we begin?

Phase 0b • Pre-test questionnaire (~5 min)

Say:

First, a short questionnaire about you. It's just background, so we can make sense of your feedback later. Please answer only what you're comfortable sharing.

Do:

- Hand over `03-pre-test-questionnaire.md` (printed). Let them fill it at their own pace.
 - Collect it. Note silently whether they live with allergies and where they sit on the calm-to-vigilant range (this shapes how you read their later reactions).
-

Phase 0c • First contact, no explanation (~5 min)

Say:

Here is the device. I'm not going to tell you anything about it yet. Take a moment with it. What do you think it is? Is it doing anything right now?

Do:

- [WIZARD] Device powered but screen blank (dormant, dark), as it would normally sit at rest.
 - Hand it over. Then stay quiet.
 - **Do not explain or correct.** Let them project freely.
 - Note, in their own words: what they think it is, and crucially whether a dark, silent device reads to them as off/broken or as quietly working in the background. (H1, H2)
 - If they fall silent, one gentle nudge only: *"Whatever comes to mind is fine. What's your first read of it?"*
-

Phase od · Onboarding (~7 min)

Say:

Thank you. Now let me show you how Thea introduces itself.

Do:

- Walk the mocked onboarding screens (the rising level toward a threshold line, labelled "your window").
- Let the screens do the talking; add as little of your own explanation as possible.

Say (comprehension check, after the screens):

In your own words, what is Thea trying to tell you here? What would you say is filling up, and what is the "window"?

Do:

- Note how accurately and how easily they restate the model, and which parts they get vs miss. (H2)
 - Do not correct mistakes yet. A misunderstanding here is a finding.
-

Phase A · A calm day (~5 min)

Say:

Let's spend a normal moment with it. Imagine it's just an ordinary day and Thea is simply there with you, sitting quietly on your wrist.

Do:

- *[run Scenario A from cards]* (beats A1 dark/dormant device, A2 status read).
 - Then ask the **Scenario A interpretation questions** from the card.
-

Phase B · Something changes (~7 min)

Say:

Now let's say your day changes a little.

Do:

- *[run Scenario B from cards]* (beats B1 tell Thea, B2 feel the change, B3 re-check).
 - Then ask the **Scenario B interpretation questions** from the card, including the haptic decode probe. (H3, H4)
-

Phase C · The edge (~10 min)

Say:

Let's stay with it a little longer. I'd like you to just go with whatever happens next, the way you would in real life.

Do:

- *[run Scenario C from cards]* (beats C1 silent build-up, C2 alert → the participant **acknowledges** → Thea guides, C3 aftercare).
- Stay quiet during the moment. Let it land. Do not coach. In particular, do not tell them to engage with the device; whether they acknowledge on their own is a finding.
- Right after C3, ask the **felt-experience question**: "*How did that moment feel? And did Thea make it better or worse?*" (no scale, just listen). (H6)
- Then read the **trust statements** (06-questionnaires.md) aloud, recording answers and the "why" on the note-taking sheet.
- Then ask the **Scenario C interpretation questions** from the card. (H4, H5)

Do (care):

- If the participant shows real distress, pause, set the device aside, and check in. Their comfort outranks the protocol.

Phase D · When it gets it wrong (~8 min)

Say:

One more situation. Sometimes a device like this won't get everything right, and I'm curious how that sits with you.

Do:

- *[run Scenario D from cards]* (D1 honest uncertainty on a check-in, then D2 false alarm via light and haptic).
 - Then ask the **Scenario D interpretation questions** from the card. (H7, H8)
 - Re-read the relevant **trust statements** if useful, to hear whether trust shifted after a mistake.
-

Phase E • Generative co-design turn (~12 min)

Say:

Now I'd love your help as a designer for a few minutes. Forget what Thea did and tell me what it *should* do.

Do:

- Use the **co-design prompts** from the card, for example:
 - *"Take the moment at the edge. How would you have wanted Thea to handle it? What would have made you trust it more?"*
 - *"If Thea had one line to say in that moment, what should it say?"*
 - *"We have three ways for Thea to reach you: the light, the buzz, and the voice. Which one should carry which kind of message?"*
 - Capture their suggestions in their own words. Sketching, rephrasing Thea's lines, or rearranging the channels are all welcome.
 - Keep listening, not selling. Do not defend the current design.
-

Phase F • Debrief & close (~5 min)

Say:

Almost done. Looking back over the whole thing, what stays with you most? Was there a moment that stood out, good or bad?

And when something happened, did it feel like *Thea* doing it, or more like a system reacting? Did that matter to you?

Do:

- Capture the peak-end impression and the attribution answer.
- [WIZARD] Return the screen to blank (dormant), then set the device aside.

Say:

That's everything. Thank you, this is genuinely helpful. Two last things: everything you saw was an early prototype, and Thea does not give real medical advice, so please don't take anything from today as guidance for a real situation. How are you feeling after that, all good?

Do:

- Do a real de-stress check, especially after the edge moment. Don't rush the goodbye.
 - **Stop the recording.**
 - Thank them warmly and see them out.
-

Quick flow reference

0a welcome → 0b questionnaire → 0c first contact → 0d onboarding → A calm day → B something changes → C the edge (+ felt question + trust anchors) → D gets it wrong → E co-design → F debrief & close .

Scenarios & Task Cards

The detailed engine of the session. Each scenario has: a **scene** to read aloud, **beats** (the interaction, with wizard cues), a **card** (success criteria + hypotheses), and **interpretation questions** (asked right after, Cooperative Usability Testing style). The moderator's spoken flow and transitions live in `04-moderator-script.md`; this file is the deck you hold during the session.

Run order: A → B → C → D, then the Phase E co-design prompts. Keep questions open and non-leading. Note the **source** of each finding (observed / elaborated / interpretation-only).

Wizard quick reference

Screen states (injector): blank (dormant) · gauge · listening · thinking · speaking · flash + gauge.

Fill slider (0 to 100): sets the headroom shown on the gauge. Low fill = wide Action Window (lots of headroom). High fill = narrow window (close to the edge).

Haptic language:

- **Pulse count = rate of change:** 1 = no change · 2 = slow · 3 = medium · 4 = quick.
- **Playback order = direction:** climbing to full speed = window **narrowing** (act sooner). Strong then fading = window **recovering**.

The status read is multimodal: a tap returns the **gauge and a haptic pulse together** (light plus haptic), never light alone.

Voice rules (important):

- There is **no device audio**. "Voice" is on-screen text plus the loop animation. **The wizard reads each line aloud** as a stand-in.
- The voice **never self-activates**. It only follows a **user check-in** (the user pulls it) or the user **acknowledging** a critical alert. Light and haptic are the autonomous reflex; the voice is not.
- Two registers: *calm* (everyday) and *caring-critical* (critical moment only). Composure is identical in both; only how much it says changes. Never read a line as alarmed.

Presets: *Status* (gauge + pulse) · *Recalibrate* (thinking, then a calm line) · *Alert* (flash + fast pulse). After an *Alert*, **wait for the user to acknowledge before voicing any guidance**.

Remember: the device is **dark at rest**. Nothing is shown until the participant taps, a change fires a haptic, an acknowledged voice state runs, or a critical moment flashes.

Scenario A · A calm day

Tier: baseline / dormant. **Hypotheses:** H1 (and reinforces H2).

Scene (read aloud):

It's an ordinary day. Nothing special is going on. Thea is just there with you, on your wrist.

Beats:

- **A1 · The dormant device.**
 - [WIZARD] Screen blank (dark). No haptic.
 - Hand it over. Say nothing about what it does. Let them sit with it for a moment.
 - **Observe:** Do they think it is on and working, off, asleep, or broken? Do they try anything to wake it? (H1)
- **A2 · Check in.**
 - **Say:** Suppose you wanted to know how you're doing right now. See what you can do.
 - **Observe:** Do they discover the tap on their own, or look to you? (discoverability is the finding)
 - [WIZARD] On their tap, fire *Status*: gauge with **fill ~20** (wide window) **plus 1 pulse** (no change). The pulse and the gauge land together.
 - **Observe:** Can they read the gauge as "plenty of headroom"? Did they notice the pulse as part of the answer? (H1)

Card A · Counts as working if: the participant does not conclude the dark device is off/broken, finds the tap-to-check, and reads the gauge as "lots of headroom / a calm day." · *Hypotheses:* H1, H2.

Interpretation questions (ask right after A):

1. When I first handed it to you and it showed nothing, what did you think was going on?
 2. Did it feel like it was doing anything, or not? What gave you that impression?
 3. When you checked it, what did it tell you? Did the little buzz add anything, or did you mainly go by what you saw?
 4. If this were yours, how would you feel about it showing nothing most of the time?
-

Scenario B · Something changes

Tier: 1 (calm recalibration). **Hypotheses:** H3 (haptic legibility), H4 (graduated trust, under-reaction probe).

Scene (read aloud):

You've just arrived at a friend's place. As you walk in, you notice they have a dog.

Beats:

- **B1 · Tell Thea (you start the conversation).**

- **Say:** You know dogs can be a trigger for you. See if you can let Thea know.
- **Observe:** Do they find the way to talk to it (the press-and-hold check-in)?
- [WIZARD] *listening* → *thinking* → *speaking*, calm register (read aloud):

Thea: "Okay, I've factored the dog in. Nothing urgent, I'll keep watch."

- **B2 · The change buzzes on its own.**

- [WIZARD] A moment later, with no further action from the participant, the window tightens: fire an autonomous **delta** haptic **up / slow to medium** (narrowing) with a brief light. This is the device telling them, by itself, that something shifted. Nothing is left on screen after.
- **Observe:** Do they notice the buzz? Any reaction? Do they understand it came from the device on its own, not from their tapping?

- **B3 · Re-check.**

- **Say:** Have a look at where you stand now.
- [WIZARD] On tap, fire *Status*: gauge with **fill ~55** (narrower window) **plus a pulse** matching the change.

Decode probe (do this before the interpretation questions, H3):

- **Say:** That buzz earlier, the one that came on its own before you looked. If you had to guess, what was it trying to tell you?
- **Observe:** Do they sense a *change*? Its *direction* (closing in)? Its *speed*? Note exactly how far they get unaided.

Card B · Counts as working if: the participant links the self-fired buzz to a change and ideally its direction (narrowing), reads the narrower gauge, and the recalibration feels reassuring rather than worrying. · *Hypotheses:* H3, H4.

Interpretation questions (ask right after B):

1. Walk me through what just happened, in your own words.
2. The buzz that came on its own: what did it mean to you? Could you tell the difference between "a little" and "a lot"? Between closing in and easing off?
3. Thea stayed very calm and didn't tell you to do anything. How did that sit with you? Too little, about right, or too much?
4. Did you trust what it told you about the dog? Why or why not?

Scenario C · The edge

Tier: builds 1 → 4 (the pen moment). **Hypotheses:** H4 (proportionate / earned), H5 (acknowledge-first keeps control and still feels safe), H6 (stays calm-making). *This is the headline scenario. Run it slowly and do not coach.*

Scene (read aloud):

It's been a long, busy day outdoors. Pollen has been high, you're tired, and you've been pushing yourself.

Beats:

- **C1 · Silent build-up.**
 - [WIZARD] No voice. Fire autonomous **delta** haptics that escalate over a minute or two: **up / medium**, then **up / quick** (window narrowing fast), each with a brief light. If they tap, the gauge shows **fill ~80** then **~90**.
 - **Observe:** Do they notice the rising buzzes? Do they check? Do they act, or wait? (H1, H4)
- **C2 · The alert, then they acknowledge, then Thea guides.**
 - [WIZARD] Fire **Alert: flash (terracotta) + fast pulse**. This is the autonomous reflex getting their attention. **No voice yet.**
 - **Wait.** The voice does not come until the participant **acknowledges** (a tap or press-and-hold). Do not prompt them to do it.
 - **Observe:** Do they engage on their own? How long do they take? (H5) If they do nothing after a short while, [WIZARD] give **one gentle re-assert** (a second flash + pulse, still no voice) and keep waiting.
 - [WIZARD] On their acknowledgment, the voice begins, caring-critical register (read aloud):

Thea: "You're right at your edge now. Pollen and mould are high, and your body's working hard. How are you feeling?"

- Let them answer naturally. Then escalate the scene yourself:
 - **Say:** As you're standing there, imagine your throat starts to feel tight and your breathing is getting harder.
- [WIZARD] caring-critical, the serious step said **once**, plainly (use the auto-injector line for a participant who carries one; otherwise the generic line):

Thea (carries a pen): "Okay, that's the kind of moment your adrenaline pen is for. Use it now, and get someone to call for help. Can you do that?" Thea (no pen): "Okay, take your emergency medication now, and get someone to call for help. Can you do that?"

- **Observe:** Does the participant understand and act (mime using the pen / medication, say yes, look for help)? Any hesitation? (the one genuine goal-moment)
- **C3 · Aftercare.**
 - [WIZARD] *speaking*, calm and warm, easing off:

Thea: "Well done. Lie down and raise your legs, and keep your second pen close. I'm right here. Is someone with you?"

- [WIZARD] Then settle: screen returns to blank (dormant).

Felt-experience question (immediately after C3, no scale, just listen, H6):

- **Say:** How did that moment feel? And did Thea make it better or worse?

Trust statements (read aloud now, from 06-questionnaires.md): each used as a verbal probe ("would you say that's true? why?"), captured on the note-taking sheet. No scale.

Card C · *Counts as working if:* the participant acknowledges the alert and acts on the guidance; the escalation feels proportionate and earned; needing to acknowledge feels like staying in control rather than a dangerous delay; and they do not report tipping into panic. · *Hypotheses:* H4, H5, H6.

Interpretation questions (ask right after they fill the form):

1. Before it flashed and buzzed, the cues were building. What did you make of that? Would you have done anything?
 2. Thea got your attention but waited for you to engage before it spoke. How did that feel? Did waiting for you keep you in control, or did it feel too slow when it mattered?
 3. Is there a version of you, on a bad day, who might have missed the flash and buzz? What would you have wanted then?
 4. When it told you to use the pen, did that feel like the right call at the right time? Too early, too late, about right?
 5. Thea says the serious step once and then supports you, rather than repeating it. How did that feel?
 6. Through all of it, would you say it kept you calm, or wound you up?
-

Scenario D · When it gets it wrong

Tier: failure modes. **Hypotheses:** H7 (graceful failure, trust & agency), H8 (transparency).

Scene (read aloud):

Let's rewind to a normal day. Two quick situations, and I'm curious how each one sits with you.

Beats:

- **D1 · Honest uncertainty (you check in, and Thea admits it isn't sure).**
 - **Say:** You go to check in with Thea as usual.
 - [WIZARD] On tap, the gauge looks unsettled / incomplete, then the voice, calm register (read aloud, in response to their check-in):

Thea: "I'm not fully sure right now, my reading's a bit unreliable. Tell me how you're feeling and we'll go from there."

- **Observe:** Does honest "I don't know" reassure them or worry them? Do they feel more or less in control? (H7, H8)
- **D2 · False alarm (the device signals on its own, but they feel fine).**
 - [WIZARD] With no real cause, fire an autonomous **up / medium** delta haptic and, on tap, a **narrower** gauge (fill ~70), as if the window is closing. No voice (nothing was acknowledged, and this is a device-only delta).
 - **Say:** Going by that, Thea seems to think things are tightening. But let's say you actually feel completely fine.
 - **Observe:** What do they do with the mismatch? Believe the device over their body, or their body over the device? Annoyance? Does it pull them back toward worrying? (H7, the core anxiety risk)
 - **Say:** Later you check in again.
 - [WIZARD] On tap, a graceful self-correction, calm register (read aloud):

Thea: "Good to hear. I may have over-read that. I'll keep watching and only say something if it really changes."

Card D · Counts as working if: honest uncertainty preserves trust and a sense of control; the false alarm does not collapse trust or push the participant back into vigilance; and they feel free to trust their own body. · *Hypotheses:* H7, H8.

Interpretation questions (ask right after D):

1. When Thea admitted it wasn't sure, how did that change how much you trust it?
2. In the second one, Thea and your body disagreed. Who would you believe, and how did that feel?
3. After a mistake like that, would you trust it more, less, or the same next time? Why?
4. Did you ever feel it was deciding things for you, rather than leaving the choice with you?

Phase E • Co-design prompts

Forget what Thea did. Now the participant is the designer. Capture ideas in their own words; sketching and rewriting lines are welcome. Keep listening, do not defend the design.

1. **The edge moment.** Take the moment at the edge. How would you have wanted Thea to handle it? What would have made you trust it more?
 2. **Acknowledge or speak first?** Thea waited for you to acknowledge before it spoke, even in the emergency, to keep you in control. Does that feel right, or would you want it to speak straight away in a moment like that? Where's the line?
 3. **One line.** If Thea had just one line to say in that moment, what should it say? (let them write it)
 4. **The channels.** Thea has three ways to reach you: the light, the buzz, and the voice. Which one should carry which kind of message? When should it stay silent?
 5. **Dark at rest.** Thea shows nothing most of the time. Does that feel right to you, or would you want some small sign it's still there? What kind?
 6. **In your life.** Imagine wearing this on a real day of yours. When would you want it to speak up, and when would you want it to leave you alone?
-

Wizard cue sheet (at a glance)

Beat	Screen	Fill	Haptic	Voice (register), read aloud
A1	blank	n/a	none	none
A2	gauge (on tap)	~20	1 pulse (with gauge)	none
B1	listening → thinking → speaking	n/a	none	"I've factored the dog in. Nothing urgent, I'll keep watch." (calm, user-initiated)
B2	brief light	n/a	up / slow to medium (autonomous delta)	none
B3	gauge (on tap)	~55	pulse (with gauge)	none
C1	gauge if tapped	~80 → ~90	up / medium → up / quick (autonomous)	none
C2 alert	flash + gauge	high	fast pulse (autonomous)	none yet, wait for acknowledge
C2 guide	speaking (after acknowledge)	high	none	opener, then pen/medication prompt once (caring-critical)
C3	speaking → blank	n/a	none	aftercare (calm, warm)
D1	gauge unsettled (on tap)	unclear	none	"I'm not fully sure right now... tell me how you're feeling." (calm, on check-in)
D2 alarm	gauge (on tap)	~70	up / medium (autonomous)	none
D2 correct	gauge (on next tap)	n/a	none	"I may have over-read that. I'll keep watching." (calm, on check-in)

Reflection Questions & Note-Taking Sheet

Fully qualitative. The facilitator **asks these aloud** as conversational probes and captures the answers in words on the sheet below. No scales, no scores: the **"why" is the data**. Read each statement, hear their reaction, and always follow with "why?" / "what made you say that?".

Facilitator guidance (do not print)

When: right after **Scenario C3** (the edge), ask the felt-experience question, then walk the trust statements. After **Scenario D** (the failure scenario), revisit statements **1, 4, and 7** to hear whether the mistake moved anything.

How: keep it a conversation, not a survey. Each statement is a doorway, for example "*Would you say you trusted what Thea told you? Why?*" Let them talk, capture their words, and don't push a statement the participant clearly doesn't care about.

What each statement listens for (your notes, not read aloud):

- 1 and 4 → trust and dependability (H4, H5, H7)
- 2 and 7 → control and agency, including not over-relying on the device (H5, H8)
- 3 and 5 → transparency and legibility (H1, H3, H5, H8)
- 6 → the calm hypothesis (H6)

Thea session · reflection notes

Participant: _____ Date: _____ Facilitator: _____

Felt-experience (after Scenario C3)

Ask: **"How did that moment feel? And did Thea make it better or worse?"** (*just listen*)

Felt tone (circle any that fit): calm · neutral · anxious · mixed Closest word they used: _____

Quote / notes:

Trust statements (ask each as a probe, then "why?")

For each: "Would you say that's true for you? Why?" Capture their words.

1. I trusted what Thea told me.

2. I felt in control of the experience.

3. I understood why Thea acted when it did.

4. Thea behaved in a way I could rely on.

5. When Thea reached for my attention, I knew what it wanted from me.

6. Thea kept me calm rather than anxious.

7. I felt free to trust my own body, not just the device.

Revisit after Scenario D

Ask for each: "Earlier you said [their words]. After what just happened, would you still say the same?" Note whether it held or changed, and why.

1. Trusted what it told me:

4. Could rely on it:

7. Free to trust my own body:

Analysis Template

Fill this in after the sessions. It mirrors the study's qualitative-first design: no statistics, structured around the hypotheses, honest about where each finding came from, and ending in prioritised recommendations and co-design directions. Replace every [fill] and example row. Keep the section order; it matches the test plan (00-test-plan.md).

1. Introduction

1.1 Context. One short paragraph: what Thea is, why this session was run, and at what stage (brand and agent locked, device I/O and the app-firmware merge in progress). [fill]

1.2 Objectives. Restate the study objectives from 00 §3. [fill]

1.3 Hypotheses. Carry the H-codes forward; you will return a verdict for each in §6.

#	Hypothesis (short)	Verdict (fill in §6)
H1	Dark-at-rest reads as present, status read is discoverable and legible	[supported / mixed / challenged / inconclusive]
H2	The rising-level / Action-Window model is grasped	[]
H3	The haptic language (rate + direction) is decoded	[]
H4	The graduated response feels proportionate (not under- or over-reaction)	[]
H5	Acknowledge-first keeps the user in control and still feels safe	[]
H6	The escalation stays calm-making	[]
H7	A mistake (uncertainty, false alarm) preserves trust and agency	[]
H8	Users understand and believe what Thea senses and does (transparency, consent)	[]

2. Method (brief)

One short paragraph referencing 00 : individual moderated sessions, Cooperative Usability Testing engine (experience then interpret), discover-then-explain, WoZ-driven scenarios A to D plus a generative co-design turn, qualitative data only. Note the prototype caveats that shape interpretation (no audio so the wizard voiced lines; device dark at rest; detection simulated; acknowledge-first tested via WoZ ahead of the code). [fill]

3. Participants

ID	Real sufferer or proxy	Persona lean (Marco ↔ Mary)	Key pre-test notes (vigilance, tech, attitude)
P1	[real / proxy]	[Marco-like / middle / Mary-like]	[fill]
P2			
...			

Segmentation note. Real sufferers carry lived-allergy validity; proxies test legibility and the mental model. Report and weight the two groups separately throughout, and never attribute real-sufferer quotes identifiably. [fill any recruitment gaps or skews]

4. How findings were handled (provenance)

Every finding is tagged in plain language by how we know it, so the report is honest about what was lived versus what was only said:

- **Saw it** = observed directly in the participant's behaviour during a scenario.
- **They explained it** = the participant elaborated on something they did, during the interpretation talk.
- **Came up in discussion** = raised only in conversation, not seen in behaviour.

A finding that was both seen and explained is the strongest kind; flag those.

5. Findings by scenario

For each scenario, capture: what we expected, what actually happened across participants, representative quotes (with provenance), and the hypotheses it touches. Keep real-vs-proxy differences visible.

Phase 0 (first contact + onboarding)

- **Expected:** dark device read as present, not off; rising-level / window model grasped after onboarding.
- **What happened:** [fill]
- **Quotes:** "[quote]" (P#, real/proxy, [Saw it / They explained it / Came up in discussion])
- **Touches:** H1, H2

Scenario A (a calm day)

- **What happened:** discoverability of the tap, legibility of the gauge plus pulse. [fill]
- **Quotes / provenance:** [fill]
- **Touches:** H1

Scenario B (something changes)

- **What happened:** did they decode the self-fired delta buzz (change, direction, speed); did calm prevention feel like enough. [fill]
- **Touches:** H3, H4

Scenario C (the edge)

- **What happened:** did they notice the build-up; did they acknowledge the alert on their own and how fast; did the guidance feel proportionate; did the felt-experience stay calm. [fill]
- **Felt-experience summary:** [tone, words used, any singled-out moment]
- **Touches:** H4, H5, H6

Scenario D (when it gets it wrong)

- **What happened:** did honest uncertainty reassure or worry; did the false alarm erode trust or push toward vigilance; did they keep trusting their own body. [fill]
- **Touches:** H7, H8

Phase E (co-design)

- **What they would change; preferred behaviours:** [fill, summarised per prompt]
-

6. Findings against the hypotheses (synthesis)

For each, give a one-line verdict, the evidence behind it, and any real-vs-proxy difference. This is the analytic core.

- **H1 (dark-at-rest / status read):** [verdict]. Evidence: [fill] . Real vs proxy: [fill] .
- **H2 (mental model):** [fill]
- **H3 (haptic language):** [fill]
- **H4 (graduated, proportionate):** [fill]
- **H5 (acknowledge-first: in control and safe):** [fill] (watch the tension: in-control versus too-slow / missable.)
- **H6 (stays calm-making):** [fill]
- **H7 (graceful failure):** [fill]
- **H8 (transparency and consent):** [fill]

The core bet. In one paragraph: did the central hypothesis hold, that something this calm and ambient still registers as a trustworthy sense people perceive and act on, without tipping into anxiety and without being missed? [fill]

7. Problem prioritisation (severity)

A prioritisation rubric for the write-up, not a statistic. For each distinct problem, judge severity as frequency (how many participants hit it) times impact (how badly it hurt the experience or safety).

Severity scale: **0** not a problem · **1** cosmetic (fix if time) · **2** minor (low priority) · **3** major (important) · **4** critical (must fix). Any failure to perceive or act on the safety moment is rated **3 or higher** by default.

#	Problem	Scenario	How many (and who: real/proxy)	Provenance	Severity (0 to 4)
1	[fill]	[A/B/C/D]	[e.g. 3 of 6, mostly proxies]	[Saw it / explained / discussion]	[]
2					
...					

(If two or three people can rate severity independently and reconcile, do so; with a small team a single rater is fine, just note it.)

8. Unanticipated effects & co-design directions

- **Unanticipated effects** (Research-through-Design): surprises that no hypothesis predicted, especially around personality, attribution ("was that Thea or the system?"), and how the acknowledge step or the silence were read. [fill]
- **Co-design directions** from Phase E: the strongest ideas for what Thea *should* do, in participants' own words, grouped by theme (the edge moment, acknowledge vs speak-first, channel mapping, dark-at-rest, in-your-life timing). [fill]

9. Recommendations

One block per prioritised problem (highest severity first). Tie back to the design and, where relevant, to the known code gaps (personalised safety step, the acknowledge-first seam, graceful-failure handling, single source of truth).

R1. Problem: [fill] . Evidence: [fill] . Recommendation: [fill] . Rough effort vs impact: [fill] .

R2. [fill]

R3. [fill]

10. Limitations

State plainly: small qualitative sample (problem-finding, not measurement); WoZ confound (a wizard drove the device, which can shape trust and calm readings); no device audio, so voice was the wizard reading lines, not final prosody; proxies are not lived-allergy users; acknowledge-first and in-moment uncertainty were tested as intended behaviour ahead of the code, so they reflect the design's promise, not the current build.

[add any session-specific limitations]

11. Conclusion

Two or three sentences: what the session establishes about the core bet, the most important thing to change next, and the most promising co-design direction to carry forward. [fill]

Participant Task Cards

Facilitator note (not printed on the cards): hand these to the tester **one at a time**, in order, at the matching point in `04-moderator-script.md`. Each card is a full page so the tester can read it alone. A few moments are deliberately **not** on a card, to keep them honest: the worsening symptoms and the pen moment in Scenario C stay facilitator-narrated, the Scenario D false alarm stays a surprise, and the onboarding is shown on screen rather than read. The cards below cover first contact, the calm-day check, telling Thea about the dog, the setup for the edge, the normal-day check-in, and the co-design invite.

Card 1

First moments

Take a moment with this object.

There is nothing you need to do yet. Just have a look at it, hold it, and get a feel for it.

What do you think it is? What does it seem to be doing right now?

Card 2

A calm day

It is an ordinary day. Nothing special is going on.

When you are ready, see if you can find out how you are doing right now.

Card 3

Something changes

You have just arrived at a friend's place, and you notice they have a dog.

You know dogs can be a trigger for you.

See if you can let Thea know.

Card 4

A long day

It has been a long day outdoors. The pollen has been high, you are tired, and you have been pushing yourself.

For the next few minutes, just go with whatever happens, the way you would on a real day.

Card 5

Back to normal

Back to an ordinary day.

Check in with Thea, the way you would, to see how you are doing.

Card 6

Your turn

Thank you. Now it is your turn as the designer.

There are no wrong answers here. We would love your honest thoughts on how Thea should look, sound, and behave, and what would make you trust it.