

Usability test plan

a) User population

The target user population for this study corresponds to the audience segments identified in the project proposal. Luma is designed for individuals who seek a calm, distraction-free way to begin their mornings by replacing the overstimulation caused by smartphones.

User segments:

The Pressed Scheduler

- Highly time constrained individuals such as young professionals, graduate students or parents with tight morning schedules
- Ages 18-40
- Medium to high digital literacy
- Comfortable with digital devices but prioritize efficiency
- Multitasking is part of their morning routine
- Relevance to testing: Could test how well the mirror communicates time sensitive data at a glance.

The Digital Native

- Younger users who grew up with screens and connected to technology, typically early adopters of smart home devices
- Ages 18-30
- High digital fluency
- Extremely comfortable with remote interfaces, apps and customization, likely to explore non obvious features
- They use multiple digital devices in their morning routings
- Relevance to testing: Could test the system's intuitive discoverability and appeal to those expecting sleek, responsive tech interactions

The Traditional Routinist

- Users with established morning habits who are less likely to use new technology unless it feels familiar and unobtrusive
- Ages 35-65
- Moderate to low technological confidence
- Comfortable with basic electronics but no accustomed to smart home interfaces

- Simple and repetitive morning routines, prioritize comfort and predictability
- Relevance testing: Could help evaluate the familiarity, readability and minimalism goals

The Active Achiever

- Health conscious individuals who integrate fitness or mindfulness activities into their mornings
- Ages 25-45
- Moderate to high technical confidence
- Use wearable fitness trackers, smart scales or scheduling apps
- Structured yet mindful morning routines, often check weather and schedule before heading out
- Relevance to testing: Could provide key feedback whether the display supports quick situational awareness without disrupting focus or movement

b) Usability goals with quantitative measurements

The usability goals for Luma aim to evaluate how effectively the system integrates into users' everyday morning routines while embodying the principles of calm technology, delivering useful information passively and without distraction. Each goal is grounded in Nielsen's 10 Usability Heuristics, emphasizing simplicity, recognizability and peripheral interaction, which will guide quantitative evaluation during usability testing and heuristic review.

1. User trust

Associated Heuristic: 1. Visibility of system status.

Goals: Users must be able to verify at a glance whether the displayed information is current and updated. The system must provide immediate and visible feedback when the user performs an action using the remote.

Measurement criteria:

- 100% of users locate the "Last updated" timestamp within 5 seconds
- 90% report confidence in data accuracy during post-test survey
- 100% of evaluators should confirm that a visual cue appears within a second of input
- Fewer than 1 in 5 actions ($\leq 20\%$) should be perceived as delayed

Rationale: Users need assurance that their input was received, immediate feedback reinforces user's confidence. A lack of visible response increases frustration and uncertainty, especially when interacting through a remote, subtle but clear feedback enhances user trust without breaking the device's calm aesthetic.

2. Match between system and real world

Associated Heuristic: 2. Match between system and real world.

Goals: The interface must use symbols, labels and command logic familiar from other frequently used technologies. Users should understand the meaning of icons and actions without extra explanation.

Measurement criteria:

- At least 90% of evaluators should correctly identify the function of all icons without external explanation.
- No more than 10% of users make errors in interpreting icon functions during initial interactions

Rationale: Familiarity minimizes cognitive load and ensures new users feel comfortable immediately, aligning with calm technology principles. By using common visual metaphors and language, the system avoids forcing users to learn new conventions, supporting intuitive and stress-free use.

3. Use the mirror as a mirror

Associated Heuristic: 3. User control and freedom.

Goals: The mirror must remain fully functional as a reflective surface while displaying information, displayed elements should never obstruct key facial areas. The user should be able to easily exit or recover from any action within two steps.

Quantitative measurement:

- 90% of users successfully complete grooming related tasks while information is displayed
- No more than 10% of users report visual interference during use
- 100% of evaluators should be able to return to the home display within two button presses or less
- Reported confusion during navigation should be $\leq 10\%$ of total observed interactions.

Rationale: Luma's primary purpose is being a mirror, technology should complement not dominate, ensuring users can perform their daily routines without obstruction aligns with the goal of seamless integration into their daily life. As well, maintaining a sense of control prevents frustration, calm interaction means users never feel trapped or disoriented while using Luma.

4. Consistency and standards

Associated Heuristic: 4. Consistency and standards.

Goals: All remote buttons, icons and layout conventions must behave predictably across all contexts.

Measurements criteria:

- 100% of evaluators should find control behavior consistent across all functions tested
- Severity rating for inconsistency should not exceed 2 on the 1-5 severity scale.

Rationale: Consistency reduces the need for conscious thought, enabling calm and fluid interaction. Predictable behavior helps users integrate Luma naturally into their morning routines.

5. Aesthetic and minimalist design

Associated Heuristic: 8. Aesthetic and minimalist design.

Goals: The display should present only essential information like time, date, weather and tasks; with no redundant visual elements.

Measurement criteria:

- 90% of users identify all displayed information types within 10 seconds of first glance
- 90% of evaluators should agree that the display feels minimal and non intrusive
- Average readability ratings should be ≥ 4 across different lighting conditions
- Visual clutter comments should appear in $\leq 10\%$ of quantitative feedback

Rationale: The mirror should fade into the background when not in use, the aesthetic clarity of the interface defines the calm experience, overly complex visuals create mental noise and disrupt the ambient nature of Luma.

6. Familiarity

Associated Heuristic: 6. Recognition rather than recall.

Goal: Users should instantly recognize all interface elements and understand their meaning through familiar conventions. All key functions must be discoverable through visible icons or on screen cues, no memorization or hidden gestures required.

Measurement criteria:

- At least 90% of users correctly identify all 5 interface elements (time, date weather, tasks and settings)
- 0 verbal request for clarification regarding icon meaning
- 90% of evaluators should locate the brightness control without prior instruction
- Average discovery time for these actions should be ≤ 10 seconds

Rationale: Calm interaction depends on recognition, not memory. Users should not have to remember button combinations or prior experiences to interact successfully, especially in the morning when cognitive resources are low.

Expected outcomes

We expect Luma to perform strongly in consistency, aesthetic minimalism and recognition rather than recall, reflecting the improvements made during design evolution. Moderate issues may arise under visibility of system status and user control, where fine-tuning response timing and navigation hierarchy will likely improve performance in future iterations.

c) Detail test procedure

Test roles

- 1 Facilitator: guides the participant through the session.
- 1 Note taker: records timing, observation and comments.
- 1 Photographer: captures photos or videos of users actions for later analysis.

Required equipment

- Laptop running the computer prototype
- Timer
- Observation data sheets
- Camera for photos or videos

Prototype handling

- Refer to the provided user manual for setup.
- Verify that the prototype is preloaded and fully functional.
- Calibrate screen brightness to ambient lighting.
- Ensure the remote responds to button presses on screen.
- Reset the interface to the home screen before each new participant.
- Do not alter or debug the prototype during a test session.
- If technical issues occur, the facilitator notes the interruption and restarts the prototype.

Examiner Instructions

1. Welcome and consent

Facilitator: “Hello! Thank you for coming, my name is (*facilitator's name*) and today you will try a prototype of *Luma*, a calm smart mirror that shows the time, weather and a short task list. We are testing the design, not you, so your honest feedback helps us improve the product. Is it okay if we record a video and take a few photos for analysis? We will not share them outside the research team and will remove identifying information on request, if you prefer not to be recorded, we can still proceed with note taking only. Do you have any questions before we begin?”

2. Introduction to prototype

Facilitator: “I will briefly describe the mirror, please do not interact yet, just listen. *Luma* shows the time, weather and calendar items, you control it with this TV-style remote. For this session, please try to think aloud, say what you notice and what you expect to happen when you press buttons, if you forget to speak aloud that’s fine, just continue. I will only answer procedure questions, not help with tasks”.

Demonstrate the remote briefly: point out arrow keys, OK, Back, +/- brightness and a visible gear icon on the screen. Do not demonstrate task solutions.

3. Warm up interaction

Give the remote to the participant

Facilitator: “Before we start, take 30 seconds to look at the remote and tell us what you think the main buttons do.”

*Note participant assumptions.

4. Task sequence

For every task, read the prompt, start the timer when the participant begins interacting, stop after success or the designated time, record time and take the necessary notes. Keep neutral and do not offer help

- Task 1. Complete initial setup process

Objective: Can a first time user complete setup within 5 minutes?

Script: “Luma has not been set up, please follow the on screen instructions to complete the setup. I will observe but will not help unless you are completely stuck, tell me when you think you are finished”.

Target: ≤ 5 minutes

Observe: steps causing confusion, attempts to skip and whether they read instructions.

Follow ups: “How would you describe the setup? Which step felt unclear?”

- Task 2. View today’s priorities

Objective: Can users locate to-do list while doing a grooming action?

Script: “Imagine you are getting ready and approach the mirror and point to where you think today’s tasks are shown, continue checking your appearance while doing this”.

Target: Locate within 5 seconds. “Continue for 20 seconds while glancing at the mirror”. After 20 seconds, cover the mirror and ask them to recall tasks they saw.

Observe: glancing vs stopping, recall accuracy.

Follow ups: “Did you need to stop what you were doing to read the tasks?”

- Task 3. Check current time

Objective: Is time easy to find and trust?

Script: “Please show me where the current date and time are displayed”.

Target: ≤ 2 seconds.

Follow ups: “Was the format immediately understandable? Would you trust it over your phone?”

- Task 4. View weather

Objective: Is the weather display clear (high/low/rain chance)?

Script: “Please find today’s weather and tell me the high, low and chance of rain.”

Target: Locate and verbalize within 5 seconds (processing time measured separately).

Follow ups: “Would this change what you wear or whether you bring an umbrella?”

- Task 5. Adjust display brightness

Objective: Can users quickly change brightness?

Script: “Without help, please make the display brighter, then dimmer.”

Target: Locate control within 10 seconds.

Observe: remote +/- vs settings menu, immediate feedback and accidental presses.

Follow ups: “Was this disruptive? Would you remember how to do it later?”

- Task 6. Use the mirror while reading info

Objective: Can participants perform grooming tasks while reading displays?

Script: “Use the mirror as you normally would: check appearance, fix hair, simulate brushing teeth, while also glancing at the display. I will observe for 30-60 seconds”.

Observe: whether any information blocks reflection, multitasking ability, comfort.

Follow ups: “Did the display block any important parts of your reflection?”

- Task 7. Integrate the mirror into daily routine

Objective: Evaluate mounting/placement and perceived fit in home.

Script: “Show how you would place this mirror on a dresser or hang it on a wall. Where would you put it at home?”

Follow ups: “Would you actually install this in your bedroom/bathroom? Why or why not?”

- Task 8. Access and navigate settings menu

Objective: Discoverability and navigation of settings.

Script: “Without help, access the settings and change the temperature unit. Then exit to the home display”.

Target: Access within 20 seconds, toggle and exit within 60 seconds.

Observe: which buttons tried, search patterns, confusion, success path.

Follow ups: “How did you figure it out? What made it easy or hard?”

5. Debrief and thank you

Facilitator: “Thank you, that is all. Do you have any final thoughts? We will use this feedback to refine the prototype. If you consented to recording, we will securely store and use it for analysis only”.

Examiner Conduct (DOs and Do not)

Observer/note-taker guide

For each task record:

- Start/stop times (s)
- success (Y/N), if it was completed without help
- Errors (wrong buttons, repeated presses, menu loops).
- Hesitation (long pauses before action).
- Severity rating for any issue observed.

DOs

- Be neutral and supportive
- Observe, do not lead
- Encourage users to express what they are thinking

Do not

- Do not help during tasks
- Do not explain interface logic

- Do not provide hints
- Do not reveal the expected outcomes or time targets
- Do not judge performance. Neutral, minimal responses only.

Measurement plan

What to measure

- Task completion time (seconds)
- Number of errors (wrong button, repeated action, misclick)
- Success rate (% completed without assistance)
- Hesitations/confusion instances
- Subjective comfort rating (1–5 scale post-task)

How to measure

- Timer and manual note taking for each task
- Tick boxes for success/failure
- Short quotes or paraphrased remarks for qualitative insight

When to Measure

- During each task (objective measures)
- Immediately after each task (subjective rating)
- After all tasks (overall impressions and satisfaction)

Data Sheet template

Task	Completion time (s)	Success (□/x)	Errors	Observed hesitation (□/x)	Participant comments	Observer notes

d) Recording and Reporting instructions

Scoring and immediate post session steps

- Facilitator and note-taker compare notes and reconcile discrepancies.
- Enter data into the spreadsheet
- Tag any critical issues (severity 4–5) for immediate design review

- Save video/photos

Data collection

Each session's performance will be logged on standardized data collection sheets including,

- Task completion time
- Number of errors or hesitations
- Success/failure outcome
- Observed user behaviors
- Verbal comments (quotes or paraphrased feedback)

Multimedia recording

Short video clips or photographs should capture,

- Interaction with the remote
- Facial expressions during confusion or satisfactions
- The mirror interface in use

Reporting procedure

Calculate quantitative averages and include qualitative observations to support interpretation, compile all results into a summary table per task showing,

- Average completion time
- Success rate
- Common issues
- Suggested improvements

Task	Average completion time	Success rate	Common issues	Suggested improvements

Expectations

- Major and minor usability issues (rated by severity 1–5)
- Patterns across participants
- Proposed design improvements to enhance calm interaction and reduce cognitive load

Rubric

User population: Does the described user population required for the test session correspond to the user segments defined in the project proposal? description of user population required for the test session corresponds to the user segments defined in the project proposal

Usability goals: What usability goals will be tested? each usability goal is linked to a quantitative measurement (where appropriate)

Usability test procedure: The detailed usability testing plan should be complete and contain appropriate detail to enable an independent team to run the test and gather useful data. Specifically, it should address the following questions: How many examiners are required? What equipment will the examiners need? How should your prototype be handled? (manual/documentation) How will you instruct your examiners to proceed? How should the examiners treat the test subjects, what should they tell them? What should the examiners avoid doing? What should the examiners avoid telling the test subjects? What/How/When should the examiners measure? test procedures addresses 6 or more of these questions

Reporting: The instructions provided to the examination team should include instructions for recording and reporting the information from the testing and evaluation procedure. explains how the information should be recorded and reported, and indicates the expectations of the design team and potential implications of the testing outcomes