THE ROBOT AGE

The Robot Experience Design (RXD)

A Framework for Evaluating the Human Side of Consumer Robotics

White Paper v2.0

Mike Posso
The Robot Age

April 2026

Suggested citation: Posso, M. (2026). The Robot Experience Design (RXD): A framework for evaluating the human side of consumer robotics (White Paper v2.0). The Robot Age.

Abstract

Consumer robotics is entering its deployment era. Service robots are in hotels and restaurants, delivery robots are on sidewalks, companion and care robots are in homes, and domestic robots are expanding well beyond the vacuum. Yet the discourse around these systems remains dominated by engineering benchmarks — autonomy, task success, safety certifications — while the human experience layer remains undertheorized and unmeasured.

This paper introduces Robot Experience Design (RXD): a structured approach to evaluating how people encounter, share space with, interpret, recover from, and fit robots into their lives. RXD is organized around six dimensions — Signal Clarity, Spatial Legibility, Perceived Presence, Failure Transparency, Interaction Fit, and Recovery Design — each addressing a distinct failure mode observed across real-world consumer robot deployments. The framework is operationalized through the Robot Experience Score (RES), a rubric-based instrument practitioners can use to audit deployments, compare products, and guide design decisions.

RXD is intended for product designers, product managers, UX researchers, deployment leads, and organizational decision-makers responsible for consumer robotics — not for robotics engineers building the underlying systems. It is released under CC BY-NC 4.0 to encourage adoption, critique, and extension by the emerging community of robotics experience practitioners.

1. Introduction: The Missing Layer

For most of its history, robotics has been an engineering discipline with engineering customers. Industrial arms on factory floors, automated guided vehicles in warehouses, surgical assistants in operating rooms — these robots were deployed into environments prepared for them, operated by professionals trained to work alongside them. The humans in these settings were onboarded, credentialed, and often literally fenced off from the machine when it moved.

Consumer robotics inverts every one of those conditions. The environment is a hotel lobby, a sidewalk, a living room, a grocery store. The user is a guest, a pedestrian, a parent, a shopper —

untrained, distracted, sometimes hostile, sometimes delighted, almost never briefed. And yet the discipline we have inherited for evaluating these systems was built for the first world, not the second.

The result is a discourse that over-indexes on what a robot can do and under-indexes on what it is like to encounter one. Engineering teams can report that a delivery robot completes 97% of its routes autonomously while being unable to say whether the pedestrians it passed understood what it was, whether it was safe to approach, or what to do if it stopped in their path. These are not edge cases. They are the primary determinants of whether a consumer robot is adopted, trusted, tolerated, or returned.

This paper proposes that the human experience of consumer robotics is a distinct domain, worthy of its own vocabulary, framework, and evaluation instruments. It introduces Robot Experience Design (RXD) as a first-version attempt at that vocabulary.

2. The State of Consumer Robotics

To understand why a framework for robot experience is needed now, it helps to look at where robotics is actually being deployed — and where the frontier is shifting.

2.1 From Speculation to Deployment

For decades, consumer robotics lived primarily in science fiction and research labs. The Roomba, launched in 2002, was for a long time the dominant example most people could name. That era is closing. In the past several years, consumer-facing robots have quietly moved from novelty to operational reality across a widening set of contexts:

- Hospitality robots delivering towels to hotel rooms and food to restaurant tables.
- Sidewalk and campus delivery robots handling last-mile logistics for food and small parcels.
- Domestic robots expanding beyond floor cleaning into mopping, lawn care, pool maintenance, and pet companionship.
- Care and companion robots in eldercare settings, pediatric hospitals, and homes supporting people with cognitive conditions.

- Retail and service robots performing inventory, cleaning, and customer-facing wayfinding in large stores.
- Humanoid platforms moving from demonstration videos toward pilots in warehouses, retail floors, and, increasingly, homes.

None of these categories is at saturation. All of them are expanding. The shift is not that any one product has become ubiquitous; it is that the aggregate surface area of human-robot contact is growing faster than the discipline that studies it.

2.2 Why Consumer Robotics Is the Frontier

There is a common assumption that industrial robotics is the mature domain and consumer robotics is the emerging one. From an engineering standpoint this is largely true. From an experience standpoint it is the inverse.

Industrial and factory robotics operate in what can be called prepared environments. The workspace is mapped. The humans present are trained operators who have been onboarded to the machine's behavior, signals, and failure modes. Safety is enforced through physical separation, interlocks, certifications, and procedure. When something goes wrong, there is a playbook and a person qualified to execute it.

Consumer robotics operates in unprepared environments. The space is shared with people who did not opt in, did not read a manual, and do not know what the robot is about to do. The user base spans children, elderly people, people with disabilities, people who have never seen a robot before, and people with strong preexisting opinions about what robots are for. The environment is variable, cluttered, and full of edge cases the robot was not trained on.

This is the frontier not because the engineering is harder — in many cases, industrial robotics is more technically demanding — but because the design problem is harder. In prepared environments, the human adapts to the robot. In consumer environments, the robot must be legible to the human, or adoption fails. The determining variable is no longer task completion. It is trust, comprehension, and fit.

2.3 The Near-Term Trajectory

Three trajectories are worth naming because they shape the urgency of this work.

First, hardware costs and capabilities are moving in the same direction — down and up respectively — that drove the smartphone era. Mobile manipulators, bipedal platforms, and sensor suites that were research artifacts five years ago are now commercial products. This expands the set of organizations deploying consumer robots well beyond the firms that designed them.

Second, the professional layer around consumer robotics is forming. Product designers, UX researchers, deployment leads, change managers, regulatory specialists, and trainers are being asked to make decisions about robots without an established body of practice to draw on. They are adapting UX methods, HRI findings, and service design tools in ad hoc combinations. A shared framework is overdue.

Third, the failures are starting to matter. Delivery robots blocking wheelchair users. Hospitality robots confusing guests. Companion robots producing unintended emotional attachments or inappropriate responses. Each of these incidents is a preventable experience failure, not an engineering failure. The absence of a shared evaluation language makes each one look like an isolated anomaly rather than a pattern the field could learn from.

The remainder of this paper proposes that language.

3. Defining Robot Experience

Before introducing the framework itself, it is worth being precise about what Robot Experience is — and what it is not.

3.1 What Robot Experience Is

Robot Experience is the designed human side of a robot encounter. It is the sum of what a person perceives, understands, feels, and does in the presence of a robot — before, during, and after the interaction.

The unit of analysis is the encounter, not the robot. A single product, deployed into two different contexts, produces two different robot experiences. A delivery robot that performs well on a university campus may fail in a dense urban neighborhood not because the robot changed, but because the encounter did. Robot Experience is therefore fundamentally situational, and evaluation must account for the deployment context, not just the device.

The practitioner of Robot Experience is the person responsible for designing, evaluating, or deploying that encounter. In most organizations, this role does not yet have a clean name. It is distributed across product designers, UX researchers, product managers, service designers, and deployment leads. One of the goals of this framework is to give that distributed practice a shared vocabulary.

3.2 What Robot Experience Is Not

Robot Experience is not robot design. Robot design is concerned with the device itself — its mechanics, its autonomy stack, its industrial design. Robot Experience is concerned with the human who meets that device.

Robot Experience is not Human-Robot Interaction (HRI) in the academic sense. HRI is a rich research tradition that has produced decades of valuable findings, and this framework draws on that work. But HRI is primarily a research discipline organized around laboratory studies and peer-reviewed publication. Robot Experience is a practitioner discipline organized around shipped products, live deployments, and design decisions made under commercial constraint.

Robot Experience is not UX applied to robots. User experience as a discipline grew up around screens and predictable software surfaces. It assumes the user is looking at the product, has an intent, and is navigating a finite set of affordances. Robots break all three assumptions. The user may not be looking at the robot, may have no intent toward it at all, and may need to make a decision about a physical object moving through their space within seconds. UX methods transfer, but the assumptions do not.

Robot Experience sits between these fields and borrows from all of them. It is not, however, a simple combination. The rest of this paper argues that it has its own dimensions, failure modes, and evaluation logic.

4. The Robot Experience Design (RXD)

The Robot Experience Design (RXD) is organized around six dimensions. Each dimension corresponds to a distinct failure mode observable in real-world consumer robot deployments. The six dimensions are:

- Signal Clarity — can the person tell what the robot is, what it's doing, and what it is about to do?

- Spatial Legibility — is the robot’s presence in the shared environment interpretable and predictable?
- Perceived Presence — how do people emotionally and socially read the robot based on its form, voice, and aesthetic?
- Failure Transparency — when the robot fails, does the person understand what happened and what to do?
- Interaction Fit — does the robot match the task, the context, and the people it serves?
- Recovery Design — after a breakdown, how gracefully can the encounter return to a working state?

The ordering is deliberate. It roughly traces the arc of an encounter: perceiving the robot, sharing space with it, reading its social and emotional character, handling when things go wrong, judging whether it belongs, and returning to normal after a breakdown. The dimensions are not fully independent — a failure in Signal Clarity often produces failures downstream — but each is analytically distinct and can be evaluated on its own terms.

4.1 Signal Clarity

Definition

Signal Clarity is the degree to which a person can perceive, identify, and interpret a robot's current and imminent state. It answers three questions simultaneously: What is this? What is it doing right now? What is it about to do?

Why it matters

A robot that is legible at a glance reduces the cognitive load of every person who encounters it. A robot that is not legible forces every bystander to build their own theory of the machine in real time, usually under time pressure. The failure mode is confusion, hesitation, and — in worst cases — collision or conflict. Signal Clarity is the foundation on which every other dimension rests, because a person who cannot read the robot cannot make any further judgment about it.

Illustrative example

Savioke’s Relay robot — a featureless cylindrical tower on wheels — glides down a hotel corridor carrying towels. It has a smooth white shell, no visible status indicators, and produces no sound. A guest stepping out of their room does not know whether this delivery is for them, whether the robot is about to stop or continue, or how to get its attention. Relay performs its task reliably, but its Signal Clarity is minimal: there is no multi-channel communication of state, no advance indication of intent, and no signal vocabulary that a first-time guest can parse. The robot is functioning. The encounter is failing.

Evaluation questions

- Is the robot's category and purpose identifiable within the first few seconds of encounter?
- Are the robot's current state and imminent next action communicated through more than one channel (visual, auditory, motion)?
- Do those signals remain legible across user groups — including children, elderly people, people with visual or auditory impairments, and people unfamiliar with robots?
- Is the signal vocabulary consistent across the product's behaviors, or does the robot use ad hoc cues for different states?

4.2 Spatial Legibility

Definition

Spatial Legibility is the degree to which a robot's presence, path, and claim on shared space is interpretable to the people sharing that space. Where Signal Clarity is about the robot as an object, Spatial Legibility is about the robot as an occupant of an environment.

Why it matters

Humans navigate shared space through a dense vocabulary of micro-cues: eye contact, subtle direction changes, body orientation, pace. Most robots possess none of this vocabulary, yet they operate in spaces where humans expect it. When a robot's trajectory is unpredictable, its claim on space is ambiguous, or its pace mismatches the norm of the environment, it generates a constant low-grade friction with everyone nearby. This friction accumulates into broader rejection of the deployment.

Illustrative example

A Starship Technologies delivery robot — a six-wheeled, cooler-sized unit weighing roughly 40 kilograms — approaches a crowded campus crosswalk. It has no gaze, no turn indication, and moves at a pace noticeably slower than pedestrian flow. Students and wheelchair users reaching the crosswalk must individually decide whether to go around it, wait behind it, or assume it will yield — none of which the robot communicates. A 2021 Northern Arizona University study documented 40 dangerous near-misses between Starship robots and pedestrians in just five days of observation. Over thousands of encounters, this friction has surfaced as disability-rights objections and, in Toronto, an outright municipal ban. The engineering works. The spatial legibility does not.

Evaluation questions

- Is the robot's path and pace predictable at the scale of the environment it operates in?
- Does the robot communicate its intent to change direction, stop, or yield before executing the action?
- Does the robot's footprint and claim on space match local norms — queueing, sidewalk flow, table spacing, room entry?
- Is the robot interpretable to bystanders who are not the primary user, including those who did not opt into the deployment?

4.3 Perceived Presence

Definition

Perceived Presence is the degree to which people emotionally and socially read a robot based on its form, voice, movement, and aesthetic. It asks not what the robot communicates about its state or path, but what the robot communicates about its character. Does it feel safe or threatening? Appropriate or alienating? Does its presence match the emotional register of the environment it has been placed in?

Why it matters

Signal Clarity and Spatial Legibility address what a robot does. Perceived Presence addresses what a robot is — the impression it creates before any action occurs. A robot's voice, facial features (or their deliberate absence), movement style, size, and material finish all carry meaning. In consumer environments, that meaning lands before any interaction begins and shapes every interaction that follows. A robot that sounds menacing, looks clinical, or moves erratically will undermine trust even if it performs its task flawlessly.

Perceived Presence is also the most context-sensitive dimension. The same robot form — say, a tall humanoid with an expressive face — may be warmly received in a retail context and profoundly unsettling in a pediatric oncology ward. Evaluators must assess not just the robot's inherent aesthetic but its fit with the emotional expectations of the specific deployment environment and the specific population within it.

Illustrative example

1X's NEO is a full-sized humanoid home robot wrapped in a soft knit suit, with tendon-driven actuators designed for gentle movement and "emotive ear rings" that communicate internal state through light. Every design choice — the fabric body, the quiet motors, the deliberately non-rigid face — is an attempt to calibrate Perceived Presence for the most intimate deployment context imaginable: a family's living room. Whether NEO succeeds will depend on whether its presence reads as safe companion or uncanny intruder to the full range of household members, including children and elderly residents. Compare this to Jibo, the tabletop social robot whose expressive eye and warm voice produced such strong emotional bonds that users wept when its servers were shut down in 2019 — a Perceived Presence success that became a product-lifecycle failure. Both cases illustrate the same principle: Perceived Presence is a design lever with consequences that outlast any individual interaction.

Evaluation questions

- Does the robot's form, voice, and movement style produce an emotional reading appropriate to the deployment context?
- Does the robot signal safety and approachability to the populations it will encounter, including children, elderly people, and people with cognitive or sensory impairments?
- Is the robot's voice — including tone, pace, pitch, and affect — calibrated to the emotional register of the context?
- Does the robot avoid triggering the uncanny valley — the discomfort produced by near-human forms that do not fully read as human — in ways that are inappropriate for the context?
- Is the robot's aesthetic consistent with — rather than in tension with — the character of the space it inhabits?

- Does the robot's design calibrate the degree of anthropomorphism to what the deployment context can sustain? Does it invite attribution of human traits that the robot's actual capabilities cannot support?

4.4 Failure Transparency

Definition

Failure Transparency is the degree to which a person can understand that a robot has failed, what kind of failure has occurred, and what, if anything, they are supposed to do about it.

Why it matters

Consumer robots fail in public. The question is not whether failures will occur — they will, routinely — but whether those failures become minor, manageable events or cascading trust failures. Opaque failures teach users that the robot is unreliable and that they cannot help it recover. Transparent failures preserve the relationship, because the person understands what happened and their role in resolving it.

Illustrative example

In 2017, a Knightscope K5 security robot — a 400-pound, five-foot-tall bullet-shaped unit patrolling a Washington, D.C. office complex — rolled off a loose brick surface and fell into a decorative fountain. The robot offered no pre-failure warning, no distress signal, and no indication of what had gone wrong. Photos of the submerged K5 went viral, turning a navigation error into a global punchline. Later, after a high-profile NYPD pilot at the Times Square subway station, a K5 unit was found parked behind the glass of an abandoned retail storefront — neither patrolling nor visibly decommissioned, its operational state unreadable to the thousands of commuters passing it daily. In both cases, the failure was not catastrophic. The transparency was.

Evaluation questions

- Does the robot distinguish between operating, paused, degraded, and failed states in a way a non-expert can perceive?
- When the robot fails, is there a clear indication of whether the failure requires human action and, if so, what kind?

- Is the locus of responsibility for recovery clear — does the bystander know whether to help, wait, or summon staff?
- Does the failure mode preserve the dignity of both the robot and the humans nearby, or does it create a spectacle?

4.5 Interaction Fit

Definition

Interaction Fit is the degree to which the robot is appropriate to the task, the context, and the people it serves. It asks whether the robot belongs in this encounter at all — not only whether it performs.

Why it matters

A robot that works is not automatically a robot that fits. Fit is a judgment about whether the robot's presence is proportionate to the need, respectful of the context, and matched to the people involved. A robot that greets patients in a pediatric oncology ward requires a different fit judgment than a robot that delivers food to a hotel room. Poor fit produces a distinct failure mode: the deployment technically succeeds and is still experienced as wrong.

Illustrative example

Unitree's Go2 is a quadruped robot — a robot dog — originally designed as a consumer and research platform. Its form is agile, expressive, and visually striking, which has led to deployments ranging from trade show entertainment to campus security patrols to educational demonstrations. But a robot dog is not a neutral form. Deployed as a novelty greeter in a retail store, the Go2 may charm. Deployed as a surveillance platform on a university campus, the same robot — with its animal-like gait and sensor array — may intimidate. The engineering is identical. The interaction fit is not. Fit is a judgment about whether the robot's form, presence, and connotations are proportionate to the task and respectful of the people it serves. A quadruped platform that delights at a robotics expo can alienate in a hospital corridor.

Evaluation questions

- Is the robot's form and presence proportionate to the task? Could a simpler solution have achieved the same outcome without the perceptual weight of a robot?

- Is the robot matched to the population it serves — in scale, pace, tone, and accessibility?
- Does the robot respect the character of the context, or does it overwrite the context's existing norms?
- Are there groups for whom this robot's presence creates a disproportionate burden — for example, people with disabilities, children, or staff whose work is reshaped by it?

4.6 Recovery Design

Definition

Recovery Design is the degree to which an encounter, once broken, can be returned to a working state. Where Failure Transparency concerns how a failure is communicated, Recovery Design concerns what happens next.

Why it matters

Every consumer robot deployment accumulates a long tail of minor breakdowns: a missed delivery, an incorrect order, a stuck navigation, an awkward moment. The strength of a deployment is measured less by the absence of these events than by how gracefully it returns from them. A poorly designed recovery turns small failures into narrative ones. A well-designed recovery often strengthens the relationship.

Illustrative example

An iRobot Roomba fails to complete its cleaning route because a chair has been moved. It returns to its dock and issues a specific notification through the companion app: the room it could not finish, the likely cause, and a one-tap option to retry. The user moves the chair and taps retry. The failure has been absorbed into the ordinary rhythm of use. This is Recovery Design working as intended — and it is no accident. iRobot has iterated through two decades of consumer deployment, producing the most mature recovery system in consumer robotics: contextual error codes, room-specific reporting, and proportionate escalation. A competitor product facing the same obstruction might report only that cleaning was incomplete, turning a one-minute intervention into an hour of troubleshooting. The difference is not sensor quality. It is the design of recovery.

Evaluation questions

- Is there a clear path from failure back to working state, and is that path proportionate to the failure?
- Does the recovery preserve any context built up before the failure — ongoing tasks, relationships, preferences?
- Is the burden of recovery distributed appropriately between the robot, the user, and any staff or intermediaries?
- Does the recovery teach the user something useful for future encounters, or does it treat every failure as a one-off?

5. The Robot Experience Score (RES)

A framework gains traction when it is operational. The Robot Experience Score (RES) is the instrument through which RXD is applied — a rubric-based evaluation that produces a dimensional profile of a robot deployment across the six RXD dimensions.

5.1 Structure

RES evaluates a deployment along each of the six dimensions on a four-point scale. Each dimension is assessed against four to six rubric criteria derived from the evaluation questions in Section 4. The output is not a single score but a profile — six dimensional ratings that together describe the shape of the experience, along with a weighted aggregate where that is useful for comparison.

Four points, not five or ten, is deliberate. A four-point scale resists the gravitational pull of the middle. Evaluators must commit to a directional judgment: the deployment is clearly failing, failing in parts, mostly succeeding, or exemplary. In practice this produces more useful audits than finer-grained scales, which tend to cluster around the median.

5.2 The Four Levels

Each dimension is rated on the same four-level scale:

- Level 1 — Absent or harmful. The dimension is not addressed, or is addressed in ways that actively degrade the encounter.
- Level 2 — Present but inconsistent. Evidence of intent but the execution fails in common cases.

- Level 3 — Reliable. The dimension is handled competently across the expected range of encounters.
- Level 4 — Exemplary. The dimension is handled with sophistication, including in edge cases and across diverse user groups.

5.3 Evaluator Evidence

RES is not self-report. A valid RES evaluation is built on direct observation of the deployment in its real context, augmented by structured interviews with bystanders, users, and operational staff where available. An evaluator must be able to cite specific observed behaviors and encounters as the basis for each rating. The scoring is only as credible as the evidence behind it, which is why the full rubric (Appendix A) pairs each criterion with evidence prompts.

5.4 Reading the Profile

The primary output of an RES evaluation is the dimensional profile, not the aggregate. A deployment that scores Level 3 on five dimensions and Level 1 on Failure Transparency is telling a specific story: the robot works until it breaks, and then the experience collapses. That story is lost in a single aggregate score. Practitioners should lead with the profile and treat the aggregate as a secondary summary useful only for comparisons across many deployments.

5.5 Applied Contexts

RES is designed to be used in four primary contexts:

- Design review. Applied to a product in development, as a structured critique before launch.
- Deployment audit. Applied to a live deployment, as a health check on the human side of operations.
- Comparative analysis. Applied across multiple products in the same category, as a research or purchasing input.
- Organizational readiness. Applied to an organization considering a deployment, as a diagnostic of the fit between the robot, the context, and the operating capacity of the team.

6. Application Contexts

RXD and RES are designed to travel. A framework that only works in one kind of organization or one phase of product development has limited practical value. The following are the four primary settings in which the framework has been designed for use, and where early adopters can expect it to be most immediately applicable.

6.1 Product Design and Development

For product teams building consumer robots, RXD provides a structured lens for design critique. The six dimensions can be used as review categories for concept reviews, prototype evaluations, and pre-launch audits. A team that embeds RXD into its design process develops a shared vocabulary for debating trade-offs that would otherwise be argued as matters of taste. It also creates a continuity of evaluation — the same framework can be applied at concept, prototype, beta, and launch, producing a consistent narrative of how the experience evolved.

6.2 Deployment and Operations

For organizations deploying robots into their environments — hotels, hospitals, retailers, campuses, municipalities — RXD serves as a deployment readiness and ongoing audit tool. A deployment is not a one-time event. Robots behave differently at different times of day, different seasons, different staffing levels. An RES audit applied quarterly provides a longitudinal view of how the experience is holding up and where intervention is needed.

6.3 Research and Editorial Analysis

For researchers, journalists, and analysts covering consumer robotics, RXD provides a citable framework for comparative analysis across products and deployments. The Robot Age uses RXD internally to score and compare consumer robots across its own research outputs, and releases the framework under CC BY-NC 4.0 specifically so that other research and editorial organizations can do the same.

6.4 Practitioner Certification

RXD is the underlying framework for the Robotics Experience Practitioner (REP) certification offered by The Robot Age. Certified practitioners are trained to administer the Robot

Experience Score as part of organizational deployments, including the Robot Readiness Audit that is the program's capstone. The framework's public release is intentional: the certification exists to produce practitioners who can apply the framework credibly, not to gatekeep the framework itself.

7. Limitations and Future Work

RXD v2.0 is a second version, not a finished discipline. It was developed through synthesis of existing HRI literature, UX and service design practice, and direct observation of consumer robot deployments. It has been pressure-tested against a range of real products but has not yet been subjected to the scale of empirical validation that a mature instrument requires. The following limitations are explicitly acknowledged, and each represents an axis of future work.

7.1 Scope

RXD addresses the encounter layer of consumer robotics. It does not address several adjacent domains that matter to the full experience of a robot in the world. Long-term relational dynamics — how a person's relationship with a household robot evolves over months or years — are out of scope for this version. Cross-cultural variation in the interpretation of robot cues is acknowledged but not systematically handled; the examples in this paper skew toward North American and Western European deployment contexts. Industrial and professional robotics are explicitly excluded; the framework is built on the assumption of untrained, unprepared users in shared environments.

7.2 Measurement

The four-level RES scale is designed for practitioner use, not for statistical power. It produces audits that are directionally useful and comparable across evaluators, but it is not a validated psychometric instrument. Future work will include inter-rater reliability studies and, where appropriate, calibration against outcome measures such as adoption, abandonment, and incident rates.

7.3 Representation

Every evaluation framework encodes assumptions about whose experience counts. RXD's evaluation questions explicitly include edge populations — children, elderly people, people with disabilities, people unfamiliar with robots — but v2.0 does not include dedicated dimensions for accessibility, cultural appropriateness, or equity of access. A future version may split these concerns out as first-class dimensions rather than treating them as cross-cutting questions within the existing six.

7.4 Roadmap

RXD v3.0 will be informed by three inputs: a growing corpus of RES audits conducted by certified practitioners, a formal inter-rater reliability study, and structured feedback from adopters using the framework under CC BY-NC 4.0. Readers who apply RXD in their own work are invited to share findings, critiques, and case studies with The Robot Age for inclusion in the next version.

8. Conclusion

Consumer robotics is past the point where engineering performance alone determines success. The robots are capable enough. What is scarce — and what will determine which deployments flourish and which get rolled back — is the design of the human encounter.

Robot Experience Design is an attempt to name that work and give it structure. Signal Clarity, Spatial Legibility, Perceived Presence, Failure Transparency, Interaction Fit, and Recovery Design are not the final vocabulary of the discipline. They are a starting point, offered to a field that has been doing this work without a shared language for too long.

The framework is released under Creative Commons Attribution-NonCommercial 4.0 International. Practitioners, researchers, and organizations are invited to use it, apply it, critique it, extend it, and publish with it. The credibility of this work will be earned not by the document but by the audits, deployments, and design decisions that draw on it.

Appendix A: Robot Experience Score (RES) — Full Rubric

The following is the v2.0 rubric. Each dimension is rated on the four-level scale described in Section 5.2. For each criterion, evaluators rate from Level 1 (absent or harmful) to Level 4 (exemplary), supported by evidence from direct observation or structured interview.

A.1 Signal Clarity

Signal Clarity — Rubric Criteria	Evidence Prompt
Category and purpose identifiable within first few seconds of encounter.	*Observation of first-time encounters.*
Current state communicated through multiple channels (visual, auditory, motion).	*Inventory of state signals.*
Imminent action indicated before execution.	*Observation of state transitions.*
Signals legible across user groups including children, elderly, and users with sensory impairments.	*Structured interviews with diverse users.*
Signal vocabulary consistent across the product's range of behaviors.	*Signal audit across behavior set.*

A.2 Spatial Legibility

Spatial Legibility — Rubric Criteria	Evidence Prompt
Path and pace predictable at the scale of the operating environment.	*Timed trajectory observation.*
Directional intent communicated before changes in direction or stops.	*Observation of navigation events.*
Footprint and claim on space match local norms for the deployment context.	*Comparison to baseline human behavior.*
Interpretable to bystanders who are not primary users and did not opt in.	*Bystander interviews.*
Navigation is robust across routine edge cases of the environment.	*Extended observation across conditions.*

A.3 Perceived Presence

Perceived Presence — Rubric Criteria	Evidence Prompt
Form, voice, and movement produce an emotional reading appropriate to the deployment context.	*First-impression observation with diverse users.*

Robot signals safety and approachability to the full range of encountered populations, including children, elderly, and people with cognitive or sensory impairments.	*Structured interviews with diverse user groups.*
Voice tone, pace, pitch, and affect are calibrated to the emotional register of the context.	*Voice audit across deployment scenarios.*
Robot avoids triggering the uncanny valley in ways inappropriate for the deployment context.	*User reaction observation; affective response survey.*
Robot's aesthetic is consistent with the character of the deployment space rather than in tension with it.	*Context audit; stakeholder interviews.*
Degree of anthropomorphism is calibrated to what the deployment context can sustain; robot does not invite attribution of human traits its capabilities cannot support.	*Longitudinal user attachment review; expectation gap analysis.*

A.4 Failure Transparency

Failure Transparency — Rubric Criteria	Evidence Prompt
Operating, paused, degraded, and failed states are distinguishable by non-experts.	*State inventory with user testing.*
Failure mode indicates whether human action is required.	*Observation of induced failures.*
Locus of responsibility for recovery is clear to the person nearest the robot.	*Bystander and staff interviews.*
Failure mode preserves dignity of robot and nearby humans; does not create spectacle.	*Public failure observation.*
Failure communication is proportionate to the severity and urgency of the failure.	*Failure taxonomy review.*

A.5 Interaction Fit

Interaction Fit — Rubric Criteria	Evidence Prompt
Form and presence are proportionate to the task; a simpler solution would not have sufficed.	*Task-to-form analysis.*
Matched to the served population in scale, pace, tone, and accessibility.	*Demographic fit review.*

Interaction Fit — Rubric Criteria	Evidence Prompt
Respects the character and existing norms of the deployment context.	*Context audit before and after deployment.*
No user group bears a disproportionate burden from the deployment.	*Equity review including staff and bystanders.*
Presence is justified by a genuine experience or operational gain, not novelty.	*Outcome evidence review.*

A.6 Recovery Design

Recovery Design — Rubric Criteria	Evidence Prompt
Clear and proportionate path from failure back to working state.	*Recovery flow walkthrough.*
Recovery preserves relevant context — ongoing tasks, preferences, relationships.	*State persistence review.*
Recovery burden distributed appropriately across robot, user, and any staff.	*Effort audit per failure type.*
Recovery teaches the user something useful for future encounters.	*User learning review.*
Long-tail minor failures are absorbed into ordinary use rather than becoming narrative events.	*Longitudinal usage review.*

References

[1] Mori, M. (1970). The uncanny valley. *Energy, 7*(4), 33–35. Translated by MacDorman, K. F. & Kageki, N., *IEEE Robotics & Automation Magazine, 19*(2), 98–100, 2012.

[2] Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems, 42*(3–4), 167–175.

[3] Bartneck, C., Belén, T., Saez-Pons, J., & Eyssel, F. (2020). *Human–Robot Interaction: An Introduction.* Cambridge University Press.

[4] Currano, R., Park, S. Y., et al. (2021). Observed sidewalk autonomous delivery robot interactions with pedestrians and bicyclists. *Transportation Research Interdisciplinary Perspectives, 19.*

[5] Ackerman, E. (2019). Anki, Jibo, and Kuri: What we can learn from social robots that didn’t make it. *IEEE Spectrum.*

[6] Holloway, D. & Cho, J. (2024). “Competing for space”: People with mobility disabilities and autonomous delivery robots on shared pathways. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems.* ACM.

[7] Nielsen, J. (1994). 10 usability heuristics for user interface design. *Nielsen Norman Group.*

[8] International Organization for Standardization. (2014). ISO 13482:2014 — Robots and robotic devices — Safety requirements for personal care robots.

[9] International Federation of Robotics. (2025). *World Robotics 2025 — Service Robots.* IFR Statistical Department.

About the Author

Mike Posso leads The Robot Age, a research, learning, and credentialing platform at the intersection of consumer robotics and non-engineering professional practice. He is Vice President of Product Engineering at JPMorgan Chase, holds faculty appointments at the Fashion Institute of Technology and the New York Institute of Technology, and is conducting graduate research at SUNY Polytechnic Institute. His work focuses on the professional layer emerging around consumer robotics — the designers, product leaders, and deployment practitioners whose decisions will shape how robots are encountered, understood, and adopted.

About The Robot Age

The Robot Age is a research, learning, and credentialing platform for the emerging discipline of robotics experience. Its mission — Robots for All — frames robotic literacy as a professional and civic competency rather than an engineering one. The platform operates across four pillars: Research, Learn, Access, and Connect. Research is the credibility engine: original frameworks, comparative analyses, and published work that sets the terms of the conversation. Learn delivers the Robotics Experience Practitioner (REP) certification and related educational programs. Access and Connect gather the practitioners, researchers, and decision-makers shaping the field.

License and Citation

Suggested citation: Posso, M. (2026). The Robot Experience Design (RXD): A framework for evaluating the human side of consumer robotics (White Paper v2.0). The Robot Age.