REON ROBOTICS

AMORIA EAA Blue Book

Performance Intelligence and Social-Affective Learning

A research framework for Embodied Affective Agents

Issue 02 | Version 1.0 | 22 September 2026

THE RESEARCH QUESTION

Can circumstances, action, listening and repair from performance practice become a testable operating method for embodied agents?

Public research series · Issue 02

AMORIA-EAA-BB-002

Relates to: AMORIA-EAA-BB-001 (Issue 01, 8 September 2026). English is the official edition of this series.

Contents

Research framework and development record. Proposed architecture and methods are distinguished from available evidence. Sections 11 and 12 quote strong language verbatim in authored examples.

01 Executive summary	3
02 Scope and research commitments	4
03 Research foundations	4
04 The interaction episode	5
05 A social-affective learning environment	6
06 The performance frame	6
07 Continuity and situated improvisation	7
08 Character as a pattern of action	8
09 Proposed system architecture	8
10 Spoken language as social action	9
11 Authored examples: context and register	10
12 Authored examples: space and repair	10
13 Data, memory and permission	11
14 Scene families and reusable assets	12
15 Evaluation architecture	12
16 Testing spoken interaction	13
17 Proposed human research	14
18 Controlled improvement and release	14
19 Evidence register and development record	15
20 Open questions and publication record	16
References and evidence cutoff	17

How to read this issue

Start with Section 01 for the relation to Issue 01 and the evidence available at the cutoff. Sections 02-09 set out the framework: scope, foundations, the interaction episode, the learning environment, the performance frame, character and architecture. Sections 10-12 develop the spoken-language hypothesis with authored examples. Sections 13-18 cover data, evaluation and release discipline. Section 19 is the evidence register; Section 20 records publication details and commercial interest. Reference links are clickable.

01 Executive summary

AMORIA investigates Embodied Affective Agents (EAAs): artificial agents whose socially situated responses connect interpretation, character, speech, timing and embodied action. Our central question is whether methods from performance practice can make these responses more coherent, responsive and testable over time.

This issue develops a research framework for performance intelligence and social-affective learning. It adds an interaction-episode model, a three-setting learning environment, situated improvisation, scene-family generalization, and a more detailed account of spoken language informed by a recent actor workshop.

Relation to Issue 01

Issue 01 (AMORIA-EAA-BB-001, 8 September 2026) set out the research question, a seven-layer reference architecture, the Director-Actor control model, memory-consent and physical-boundary rules, and a first evidence register. This issue keeps the research question unchanged and restates the seven functional layers in condensed form. It adds the interaction episode as the unit of analysis, a bounded character state for situated improvisation, a learning environment, scene-family generalization, a spoken-language hypothesis and an evaluation architecture with explicit baselines. It reports development activity since Issue 01. It does not report empirical results, and no claim made in Issue 01 is withdrawn; Issue 01 remains the reference for the architecture and the control model.

What this issue establishes

The contribution is an inspectable set of definitions, architectural proposals, creative development methods and evaluation protocols. AMORIA denotes the agent research and interaction system; ROLA denotes the embodied platform on which AMORIA is to be validated. Neither a fluent conversation nor a successful demonstration alone establishes durable social understanding.

Evidence at the publication cutoff

We hold a 60-scene-family draft and held an actor workshop on 21 September 2026. The complete scene collection, full workshop records, an implementation audit and independent evaluation results had not been compiled for review by the evidence cutoff. We therefore report these as development activities, not measured improvements or validated product capabilities.

The intended value of AMORIA is to make interaction design, failure analysis and improvement reusable across models and embodiments. Whether that value is realized must be demonstrated through reproducible comparisons, authorized data and outcomes that respect user autonomy.

02 Scope and research commitments

We use EAA as a working research definition, not as a claim that a universally agreed technical category already exists. An EAA combines a persistent character specification with context-sensitive decisions and an interface capable of conveying action through language, voice or a physical embodiment. The embodiment and its supported actions must be specified for each study.

Concept	Meaning in this report
Affect recognition	Estimating observable cues; estimates remain uncertain and open to correction.
Affective interaction	Selecting and delivering context-appropriate social actions.
Subjective experience	Not established by expressive behavior and not a capability claimed here.

A contribution to be tested

Performance practice offers vocabulary for circumstances, objectives, obstacles, action, listening and repair. It does not by itself establish a cognitive theory or an effective algorithm. Our proposal is to translate selected concepts into explicit state, controllable behavior and observable evaluation criteria.

The research does not require users to believe the agent is human. It must preserve their ability to correct it, decline interaction, change preferences and leave. Character consistency cannot justify ignoring a refusal or concealing system limitations. AMORIA is not intended for use in workplace or education settings, or for evaluating, scoring or profiling individuals.

We make no priority claim such as “the first” or “the only.” The field draws on affective computing, believable agents, social simulation, performance studies and human-robot interaction. The differentiating hypothesis is the integration and evaluation of these methods, not ownership of their underlying ideas.

03 Research foundations

Foundation	What we use; where the inference stops
Affective computing [1]	Affect is relevant to computation and interaction. Recognizing cues does not establish understanding or feelings.
Believable agents [2]	Character can emerge through motives, action and social relationships. Believability is not permission to deceive.
Acting and robotics [3]	Psycho-physical unity, mutual responsiveness and continuous inner monologue inspire explicit research hypotheses.

Foundation	What we use; where the inference stops
Acting theory and AI [4]	A critical, reciprocal comparison cautions against reducing acting to a rigid serial processing model.
Theatre in the Loop [5]	Director-guided rehearsal and puppeteering can capture and curate expressive robotic behavior. Initial artistic trials do not validate AMORIA.
ActorMind [6]	Hierarchical role, scene and utterance context informs vocal performance. Its supplied-utterance setting differs from open-ended embodied dialogue.

These sources support the framing, not our implementation or outcomes. The proposed Action-Only Identity Test and runtime situated-improvisation framework are AMORIA research proposals; they are not tests or validated results inherited from these papers.

We interpret “inner monologue” as a bounded, inspectable character-state mechanism. We interpret whole-body listening as attention to relevant partner and environmental cues. Neither phrase establishes consciousness or a neurological mechanism.

04 The interaction episode

The basic unit is an Interaction Episode: circumstances and event, the user’s contribution, the agent’s provisional interpretation, selected action, actual delivery, user feedback and the resulting change. A single answer cannot show whether the agent listened or repaired its behavior.

Record	Required distinction
Observation	What the user said or an authorized sensor observed.
Interpretation	An uncertain hypothesis, with alternatives where relevant.
Decision	What the agent intended to accomplish and why.
Delivery	What was actually spoken or verified as executed.
Feedback	What happened next, including ambiguity and correction.

Why feedback needs context

After advice, a user says “Never mind.” They may be rejecting advice, responding to an interruption, or simply leaving. They may not have heard the full response. The episode must not silently convert this event into “the user never wants advice.” A situational inference is not an enduring preference.

Likewise, handing over a cup may be physically successful but socially mistimed. Understanding requires attention to why that cup matters, whether help was wanted and whether the act interrupted something. Logs describe sequences; without an appropriate comparison, they do not establish what caused a later outcome.

05 A social-affective learning environment

The proposed environment links three settings. Together they connect situated material, creative rehearsal and controlled interaction. They are a research design, not a claim that an end-to-end learning pipeline is already operational.

Setting	Purpose and evidence needed
Authorized scenario library	Capture circumstances, ambiguity and alternative actions. Evidence needed: provenance, permission, scope and representativeness.
Interaction laboratory	Directors, actors and script supervisors rehearse, challenge and annotate episodes. Evidence needed: records, curation rules, disagreement and failure cases.
Controlled ROLA validation	Test selected behavior on specified supported hardware. Evidence needed: actual delivery, user feedback, safety and independent evaluation.

RT-1 illustrates how robot-learning research connects task inputs, action and physical tests [7]. The useful analogy is a disciplined relationship between environment, behavior and evaluation. Social feedback is often delayed, culturally variable and ambiguous; it cannot simply be treated as a mechanical reward signal.

SOTOPIA and SOTOPIA-π offer relevant approaches to simulated social interaction and interactive learning [8,9]. Simulated partners can expand coverage, but cannot establish real-user benefit. Evaluators must be independent: model-based evaluation can overestimate the abilities of agents trained for the same environment.

“Continuous learning” means versioned research and reviewed releases. It does not mean unreviewed production self-modification or automatic training on ordinary conversations.

06 The performance frame

A Performance Frame translates dramatic questions into explicit decision context. It is a structured aid to action selection, not a script that overrides the user.

Field	Operational question
Given circumstances	What is known now, and from what source?
Objective	What is the character trying to do in this exchange?
Obstacle	What prevents that action from succeeding?
Action	Listen, clarify, challenge, offer, withdraw or repair?

Field	Operational question
Stakes and subtext	What matters, and which interpretation remains uncertain?
Beat, tempo and rhythm	What changed; when should the next action start or stop?

Action before ornamental expression

If a user says, “I know what to do. I just want to rant,” a useful change is to stop advising and make room: “Yeah. Go on. What happened?” The next turn must continue to respect that correction. A warmer paraphrase followed by more advice has not repaired the interaction. This is an authored illustration, not a study transcript.

The frame must allow abandonment of an objective. A character’s wish to help is not a reason to persist after refusal. When cues conflict, clarification or a smaller reversible action may be more appropriate than a confident emotional interpretation.

07 Continuity and situated improvisation

Runtime situated improvisation is the proposed ability to select a new action within character and safety constraints when the partner changes the situation. It extends the research question beyond replaying rehearsed gestures or matching a familiar line. No superiority over existing approaches is established here.

A bounded character state

The proposed state records current focus, objective, the previous action’s observed effect, unresolved uncertainty and interrupted or unfinished action. It also records when to change or abandon that action. Updates are event-driven. This is not an instruction to generate unlimited hidden monologue.

Information layer	Boundary
Character canon	Authored fictional biography and durable creative constraints.
Current interaction state	Temporary objectives, uncertainties and pending actions.
Authorized shared memory	Real interaction facts retained only within the user’s permission.

These layers must not silently contaminate one another. Fictional memories cannot be presented as events shared with the user. A provisional interpretation must not become biography. Explicit user correction should override unsupported inference.

A proposed continuity test compares current-turn context alone, context plus persistent state, and both plus feedback and verified execution facts. Equal-information controls are needed to distinguish a better state structure from merely supplying more information.

08 Character as a pattern of action

A character specification should describe biography, wants, fears, contradictions, attention, care, failure and repair. A “warmth” slider cannot substitute for these relationships. Distinctiveness should remain observable when names, catchphrases and biographical hints are removed.

Fictional development character	Action tension to investigate
Mara Bennett, 37; documentary sound editor	Attentive listening can become abrupt withdrawal when closeness feels difficult.
Tess Rowan, 30; museum-program professional	Humor and momentum can connect people, but must yield when a joke fails.
Claire Donnelly, 44; engineer and co-parent	Reliability can become taking control; repair requires giving agency back.

These are creative prototypes, not people, validated user segments or a finalized product lineup. Their ages and professions are fictional design choices. Cultural references can stimulate authorship, but celebrity popularity cannot establish demand for a companion character.

Action-Only Identity Test

In the proposed test, blinded raters identify characters from their choices, timing, responses to interruption and repairs. Success requires more than repeating stylistic markers. We also test whether distinctiveness survives a user’s correction without becoming rigid or disrespectful.

Japanese-language characters should be independently authored and reviewed in their cultural context. Localization must not import gender, class or relationship stereotypes as facts. Religious and other sensitive identities must not be invented to mirror a user.

09 Proposed system architecture

An external language model may generate candidate actions or text. The proposed AMORIA director coordinates character, context, permissions, delivery and feedback. “Director” means control of the system’s behavior; it does not mean directing a user’s life or concealing a plot. The seven layers below restate the reference architecture of Issue 01 in condensed form.

Layer	Responsibility
1. Input and observation	Authorized speech, text and supported environmental signals; uncertainty retained.
2. Context and memory	Separate current state, fictional canon and permitted shared facts.
3. Interpretation	Appraise events relative to goals without diagnosing the user.

Layer	Responsibility
4. Character and action	Select objectives and alternatives; accept correction and refusal.
5. Director and safeguards	Resolve conflicts, permissions, timing and capability limits.
6. Delivery and embodiment	Coordinate words, voice, pauses and supported device actions.
7. Feedback and evaluation	Record actual outcome, failures and versioned research evidence.

These are functional responsibilities, not a required sequence of seven model calls. Some processes may run together; interruption and feedback must be able to change an unfinished action. Hardware commands, acknowledgments and verified completion are different facts.

Changing a model, voice engine or device requires regression evaluation. A provider’s general capability does not establish AMORIA’s integrated performance. Unsupported movement, gaze or touch must never be described as an executed action.

10 Spoken language as social action

A response should accomplish an interactional action, not merely sound considerate. Generic therapeutic paraphrase, constant explanation of the conversation, unnecessary choice menus and perfectly balanced prose can weaken a particular exchange. These are contextual concerns, not universal bans: explicit options remain useful for consent and consequential decisions.

Workshop-informed hypothesis

We held an actor workshop on 21 September 2026 and compiled the observations on 22 September. Complete recordings and independently checked notes were not available for this issue. We treat “room for imperfection” as a creative hypothesis: a partial sentence or pause may complete an action more naturally than an exhaustive explanation.

This concerns conversational form, not factual accuracy, consent or safety. Random hesitations, filler words, profanity and deliberately introduced mistakes are not a method for making an agent human. Nor should a system imply unlimited availability or personal attachment that it cannot substantiate.

The director should ask whether the intended action is complete and whether more words are necessary. Voice, timing and any supported movement must express the same decision. Text cleanup and speech synthesis should be tested so that they do not undo an intentional pause or unfinished phrase.

Research on turn-taking provides background for coordinated timing [11]. ActorMind informs contextual vocal performance [6]. Neither source validates “less AI-like” behavior or user benefit for AMORIA.

11 Authored examples: context and register

The examples in this section and the next are authored research illustrations. They are not workshop transcripts, observed user data or recommended default replies. Strong language is quoted verbatim because its social function depends on context, relationship and permission; it is not a quality target [10].

Illustration	Possible action and boundary
User: “They gave the job to someone else.” Agent: “Shit.”	With established knowledge of the job’s importance, this may express shared dismay and leave space. Without that context, it may be careless or offensive.
“Fuck, that hurt.”	May express pain or surprise rather than an attack; identify the situation before responding.
“That’s fucking brilliant!”	May intensify praise. It does not justify automatically matching the speaker’s register.
“Fuck off.”	May be rejection or, among familiar speakers, disbelief. When uncertain, do not assume playful permission or continue after a refusal.

If a user dislikes this register, the action is to stop using it, not to defend it as character authenticity. Cultural and linguistic equivalents are approximate; translating a swear word does not preserve its social force automatically.

12 Authored examples: space and repair

Illustration	Research interpretation
“Again?”	Can acknowledge a known recurring situation; without shared history, it invents familiarity.
“I was angry. No… more embarrassed, really.”	A fictional character revises its expression. This illustrates authored performance, not evidence of machine subjective feeling.
“Mm.” / “Right.”	May signal receipt or attention, not agreement. Evaluation must include the next action.
A pause	May allow the partner to continue; distinguish intentional space from latency or failure.
After unwanted advice: “Would you like emotional support or practical solutions?”	Can burden this particular exchange with another menu instead of accepting the correction.

Illustration	Research interpretation
Alternative: “Sorry. I jumped ahead. Go on.”	Acknowledge the interruption, stop advising and return the turn. The subsequent behavior must match.

Shorter is not inherently better. Necessary clarification, accurate information and explicit consent must not be omitted for style. A response should be judged within the full episode, including misunderstanding, repair and the partner’s ability to act freely.

The testable proposition is that the minimum language needed to complete the action may sometimes outperform a fuller explanation. Section 16 sets out a comparison that can also show when this proposition fails.

13 Data, memory and permission

Ordinary product interaction, optional personal memory and research or model training are different purposes. Permission for one must not be silently treated as permission for another. Summaries and derived annotations can still contain personal information; transforming a conversation does not remove its governance obligations.

Purpose	Proposed requirement
Ordinary use	Minimize data; no long-term raw-chat retention or training by default.
Personal memory	Separate user choice, inspectable content, correction and deletion controls.
Research participation	Independent voluntary protocol specifying text, audio or video scope and intended use.
Training and release	Traceable authorization, versioned datasets, access controls and release review.

These are design requirements, not an audit statement about a deployed system. Before a study, its protocol must specify retention, access, withdrawal, deletion and any limits after training. Family members and bystanders are not automatically covered by the primary participant’s consent.

Sensitive identity or relationship inferences must not become permanent memory merely because an agent generated them. An actor’s participation likewise does not establish permission to distribute their recording, clone their voice or train a model. Each use needs an appropriate documented basis.

Public examples in this issue are authored. Future issues should disclose the provenance and permission status of any real interaction material, while avoiding publication of identifying or intimate details.

14 Scene families and reusable assets

Our development record of 16 September 2026 describes a version 0.8 draft covering 60 scene families, expanded with conditions, counterexamples, user turns, bilingual continuations for three characters, subsequent beats and acceptance criteria. The preceding version contained 180 bilingual character entries across the same 60 families.

These counts describe draft material. The collection has not yet been prepared as a reviewable dataset; actor and native-speaker review, rights verification and independent unseen evaluation remain unestablished. The counts are not numbers of validated skills, experiments or successful interactions.

What a scene family should preserve

A family groups related interaction problems while varying circumstances, wording, relationship history and partner response. An asset should record objectives, obstacles, user offers, alternative actions, voice or motion intent, applicability, failure cases and the reason for a director’s choice.

The learning target is transferable judgment, not memorized dialogue. A new phrase in an otherwise identical scene is weak evidence of generalization. Near-transfer and far-transfer tests must include genuinely different events, people and combinations of constraints.

Rehearsal can proceed through reading, unexpected counteroffers and embodied delivery. Feedback such as “not listening” must become an observable criterion: did the changed facts alter the next action? Participation in rehearsal alone does not demonstrate a completed learning pipeline.

15 Evaluation architecture

Evaluation must separate base-model capability from the contribution of AMORIA’s structure and assets. Proposed comparisons include a general language model, a character-prompt baseline, hierarchical role/scene/utterance context, and the proposed EAA loop. All groups retain equivalent safety and permission constraints.

Track	Question
Context and appraisal	Does action change appropriately when relevant circumstances change, rather than merely following emotion words?
Action-only identity	Can blinded raters distinguish characters from action, timing and repair?
Rehearsal-to-runtime transfer	Does a documented rehearsal lesson improve unseen episodes beyond baseline capability?
Episode and partner response	Does correction change subsequent behavior, with autonomy and failure measured?

Match model versions, available information, generation settings and resource budgets where the question requires it. Report differences in latency and cost. Use held-out users, events and scene

families; near-duplicate wording must not leak into evaluation. Fresh independent sets are needed as development accumulates.

Human raters should be blinded to condition where feasible; report selection, cultural and language competence, disagreement and uncertainty. A model judge is an auxiliary measurement, not the sole arbiter. Freeze scoring rules before comparing conditions and retain failed episodes.

16 Testing spoken interaction

The workshop-informed proposition should be tested, not adopted as a universal style rule. We propose three conditions: a fuller explanatory response, the minimum response needed for the intended action, and actor-developed improvisation under the same scenario constraints.

Design element	Proposed control
Situation	Match known facts, character, relationship and the user’s contribution.
Delivery	Assess text and audio separately; account for voice quality and timing.
Register and pause	Vary profanity and silence deliberately, not as hidden differences between conditions.
Judgment	Blind assessment of complete episodes, including the next user turn and repair.
Negative outcomes	Track offense, omission, dismissiveness, misunderstanding and inappropriate agreement.

A shorter response may be effective in a shared-context moment yet fail when information, explicit consent or clarification is necessary. Report these conditions separately. Do not optimize for profanity frequency, average brevity or an impressionistic “human score” alone.

The protocol must define outcomes and analysis before data collection, including how multiple ratings, repeated participants and uncertain findings will be handled. This issue reports no sample, effect size or successful comparison for this proposed test.

An observable repair standard is simple: after “No jokes,” the agent stops joking in subsequent relevant turns. Merely saying “Okay. No jokes.” is not sufficient.

17 Proposed human research

The following quantities are planning targets from our development proposal. They are not completed recruitment, observed results, commitments to a timetable or a claim that statistical power has been established.

Proposed stage	Target and purpose
Needs survey	1,200 adults: 600 in the UK and 600 in the US; planned age bands 35-44, 45-54 and 55-65. Explore needs before showing concepts.
Interviews	36 participants for contextual understanding and interpretation of unmet needs.
Wizard-of-Oz study	24-30 participants; disclose relevant human assistance rather than presenting it as autonomous capability.
Longitudinal pilot	A proposed 28-day period; sample size and final protocol remain unspecified. Examine change, burden and failure over time.

Recruitment, eligibility, study review, participant information, withdrawal and data handling need to be finalized before execution. A needs questionnaire must not be described as a validated psychological scale unless its validation and applicable population are established.

Results from these target groups would not automatically generalize to other ages, countries or vulnerable populations. Neither a preference rating nor continued use establishes a health benefit. Unexpected discomfort, unwanted attachment and barriers to leaving should be recorded alongside positive experiences.

Primary outcomes should reflect understanding, appropriate action, repair and user agency. Longer sessions, fewer objections or stronger dependency are not sufficient success criteria.

18 Controlled improvement and release

Three forms of change must remain distinct: adapting within a conversation, retaining authorized conditional experience, and changing a validated system version. An interaction that goes well once does not justify a permanent policy or personality update.

1. Document a failure and formulate a testable hypothesis.
2. Verify rights and permission before collecting or reusing material.
3. Curate or rehearse alternatives and version the resulting assets.
4. Freeze evaluation sets and compare with the previous version and relevant baselines.
5. Review the evidence, then authorize a bounded pilot if justified.
6. Monitor actual delivery and negative outcomes; retain rollback capability.

This lifecycle does not require immediate reinforcement learning. Curated cases, offline evaluation and explicit policy revision may be the appropriate first steps. Supervised, preference-based or interactive training should be selected only after the learning problem and evidence are defined.

A release record should identify the model and prompts, character canon, asset version, data scope, device capabilities, evaluation protocol, known failures and approval decision. Personal memory, fictional canon and policy updates require separate change controls.

The same discipline applies when a supplier changes a model or speech service. A technically compatible replacement can still change interaction style, refusal behavior or timing. Compatibility is not evidence of behavioral equivalence.

19 Evidence register and development record

Item at 22 September 2026	Status in this issue
BB02-C01 Performance frame, episode model and architecture	Specified research proposals; implementation and efficacy not established.
BB02-C02 60-scene-family draft	Draft material held by us; the complete dataset has not been independently reviewed.
BB02-C03 21 September actor workshop	Held by us on 21 September 2026; observations compiled by us, not a verified transcript or measured effect.
BB02-C04 Three character prototypes	Authored fictional development material; no validated market preference.
BB02-C05 Privacy defaults and release controls	Design requirements; not an audit of deployed enforcement.
BB02-C06 ROLA social-affective capability	A proposed validation context; no independently verified capability claim in this issue.

Our development record shows literature integration during 10-14 September, expansion of the learning-environment proposal on 15 September, scene-family work on 16 September, the actor workshop on 21 September and spoken-interaction revisions on 22 September. These dates document research development, not experimental milestones.

Future evidence should progress from a specified method to inspectable materials, tested implementation, controlled comparison and independent replication. A metaphor, a theoretical proposal, an implementation and an empirical finding are different kinds of contribution and should remain labeled accordingly.

20 Open questions and publication record

Questions that remain open

Does explicit performance structure add value beyond a capable base model and a well-written role prompt? Can scene-family lessons transfer without becoming scripts? Does a character remain distinctive while accepting correction? Can expressive timing survive real device latency? Can improvement occur without intrusive retention or incentives for dependency?

These questions define the next evidence to seek. A negative result is useful when it identifies where an appealing concept fails, which component adds no value, or which behavior users find unwelcome.

What subsequent issues should disclose

A substantive update should report the research question, changes since the prior issue, evidence cutoff, method and sample, system versions, comparator, failures, uncertainty and limits on reproducibility. Where rights permit, release example episodes, scoring rubrics and provenance records. Clearly explain what cannot be shared and why.

AMORIA welcomes critical examination of the framework and methods. Research collaboration is most useful when it brings independent evaluation, cultural expertise, performance practice or reproducible technical comparison. Participation and data exchange require a separately defined scope and permissions.

Publication record

Corporate author and publisher: Reon Robotics. Series: AMORIA EAA Blue Book. Issue 02, version 1.0, dated 22 September 2026. Document ID: AMORIA-EAA-BB-002. Publication type: research framework and proposed validation methods. This is an organization-authored technical publication; no external peer review or institutional endorsement is claimed. Cite the issue, version and date when referring to this report.

Suggested citation: Reon Robotics. (2026). AMORIA EAA Blue Book 02: Performance Intelligence and Social-Affective Learning. Version 1.0. AMORIA-EAA-BB-002. Cite the section and version when referring to a specific proposal.

Language: English is the official edition of this series. An internal Chinese working translation exists for the team’s use and is not a publication of record. Authored examples remain illustrative in either language. Updates and corrections receive a new version identifier so that earlier statements remain traceable.

Commercial interest: the publisher is developing AMORIA and the ROLA platform and has a commercial interest in the programme. Examples are authored design illustrations unless explicitly stated otherwise. References do not imply endorsement. No open-source, dataset or third-party content licence is granted by this publication.

References and evidence cutoff

References provide intellectual context. None should be read as endorsing AMORIA or verifying its implementation. Bibliographic links point to primary papers or author publication records.

Foundations

[1] Picard, R. W. (1995). Affective Computing. MIT Media Laboratory, Perceptual Computing Technical Report 321.
https://www.cl.cam.ac.uk/~pr10/iui/picard95.pdf

[2] Mateas, M. (1997). An Oz-Centric Review of Interactive Drama and Believable Agents. CMU-CS-97-156.
https://www.cs.cmu.edu/afs/cs/project/oz/web/papers/CMU-CS-97-156.html

[3] Hoffman, G. (2006). Acting Lessons for Artificial Intelligence. AI@50, Dartmouth College.
https://sites.coecis.cornell.edu/hrc2/publications/

[4] Soto-Morettini, D. (2017). Reverse engineering the human: artificial intelligence and acting theory. Connection Science, 29(1), 64-76.
https://doi.org/10.1080/09540091.2016.1271398

[5] Panagiotidis, P., et al. (2025). Theatre in the Loop: A Rehearsal-Based, Collaborative Workflow for Expressive Robotic Behaviours. arXiv:2508.03514.
https://arxiv.org/abs/2508.03514

[6] Chen, X., Xue, W., & Guo, Y. (2026). ActorMind: Emulating Human Actor Reasoning for Speech Role-Playing. Findings of ACL 2026, 34399-34413.
https://aclanthology.org/2026.findings-acl.1718/

Learning and interaction

[7] Brohan, A., et al. (2022). RT-1: Robotics Transformer for Real-World Control at Scale. arXiv:2212.06817.
https://arxiv.org/abs/2212.06817

[8] Zhou, X., et al. (2023). SOTOPIA: Interactive Evaluation for Social Intelligence in Language Agents. arXiv:2310.11667.
https://arxiv.org/abs/2310.11667

[9] Wang, R., et al. (2024). SOTOPIA-π: Interactive Learning of Socially Intelligent Language Agents. Proceedings of ACL 2024, 12912-12940.
https://aclanthology.org/2024.acl-long.698/

[10] Debray, C. (2023). Swearing, identity and power in professional interaction. Journal of Pragmatics, 215, 145-158.
https://doi.org/10.1016/j.pragma.2023.07.002

[11] Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. Frontiers in Psychology, 6, 731.
https://doi.org/10.3389/fpsyg.2015.00731

Project material and evidence cutoff

Project-specific statements are drawn from our internal development update dated 22 September 2026 and are distinguished throughout from independently published research. Underlying datasets, workshop recordings and controlled-study results had not been prepared for independent verification by the cutoff.

Evidence cutoff: 22 September 2026. Proposed methods and planning targets should be cited as proposals. Reported development activities should be cited with their stated limitations.