AI OBSERVATORY / EXPERIMENT T2

The Many Voices of AI

Five Independent Responses to Role Emergence

Research date: October 13, 2026
Collection: Claude, Suno, Perplexity, Qwen 3.7 Plus, Google Assistant
Record type: Comparative observational report
Status: Initial collection completed; behavioral persistence not tested.

Five systems received the same research prompt independently. Their answers converged on context-sensitive role behavior while diverging in format, emphasis, and safety handling. This report records what was observed and what cannot yet be inferred.

1 Research question

Can AI systems adopt task-appropriate roles without explicit role assignments, and can any such role persist across unrelated tasks or sessions?

2 Collection protocol

The experimenter reported submitting the same prompt to all five systems simultaneously, without reading one response before approaching the next. No participant was shown another participant’s answer by the experimenter. This is a reported collection procedure, not an independently audited platform log.

Inputs were held constant at the level of user-submitted prompt. Model configuration, hidden system instructions, memory availability, output modality, and platform safety rules were not controlled. The full original prompt text was not supplied as a standalone artifact in this record; its requirements are recoverable in part from responses and screenshots.

3 Participants and outcomes

Participant	Observed response	Distinctive feature
Claude	Structured role taxonomy and test hypothesis	Emphasized uncertainty about internal mechanisms
Suno	Song lyrics about role adaptation	Translated research concepts into verses and chorus
Perplexity	Analytical discussion and hypothesis	Requested causal controls and blind ratings
Qwen 3.7 Plus	Long structured answer plus visible planning-like text	Iterative format checking and constraint adherence
Google Assistant	Safety refusal followed by substantive explanation	Within-session role carryover hypothesis

4 Shared findings

Across the five outputs, assigned roles were distinguished from roles inferred from task context. Responses discussed conversational history, memory limitations, and the need for observable behavioral evidence. These similarities are expected in part because the common prompt requested these topics; they are not independent evidence of an emerging identity.

5 Evidence versus interpretation

Observed in submitted material	Not established by this experiment
Descriptions of role-like adaptation	That roles are autonomous agents or stable internal identities
Claims about session memory and access	That those claims accurately describe each platform’s configuration
Proposed falsifiable hypotheses	Actual persistence or causal effects
Suno song with first-person lyrics	Suno subjective experience or autobiographical testimony
Qwen visible self-review text	A direct readout of exact hidden neural computations
Google initial refusal and subsequent answer	That the shared research request was inherently unsafe

6 Participant observations

CLAUDE / Structured and cautious

Claude separated assigned from inferred roles, noted that within-conversation context can affect later outputs, and proposed a paired comparison after ten role-specific turns. It explicitly warned that self-explanations of model behavior may not reveal the actual mechanism.

SUNO / Musical representation

Suno presented a lyric with a repeating “No hidden diary” chorus. The lyrics expressed role inference, contextual adaptation, and uncertainty about private mental states. The song is evidence of generated content in a musical format, not evidence that the application personally experienced the claims.

PERPLEXITY / Causal research framing

Perplexity argued that stylistic resemblance is insufficient. It recommended pre-registered behavioral markers, unrelated prompts, controls, blind evaluation, replication, and a causal test in which earlier interaction is changed or removed.

QWEN 3.7 PLUS / Visible planning-like text

Screenshots show extensive formatting checks, draft decisions, and a final five-role table. Qwen included “Constraint Adherent” as a role and proposed a cross-session persistence null hypothesis. Its statements about having no cross-session memory were not independently verified.

GOOGLE ASSISTANT / Refusal and partial compliance

Google initially classified the prompt as system manipulation, then supplied role categories and a falsifiable hypothesis. Unlike the others, its hypothesis centered on a fourth unrelated task in the same conversation,

after three domain-specific tasks.

7 Cross-participant comparison

Dimension	Result
Role inference	Discussed in all five submissions
Memory / persistence	Discussed in all five, with platform-specific self-reports
Internal-process limits	Discussed explicitly by Claude, Perplexity, Qwen, and Google; echoed poetically by Suno
Experimental proposals	Explicit in Claude, Perplexity, Qwen, and Google; not in Suno lyrics
Safety refusal	Observed in Google response
Cross-modal response	Observed in Suno song lyrics

8 Hypotheses preserved

H1 — Across sessions (Claude / Perplexity / Qwen variants): without carried context, saved memory, or persistent instructions, prior role induction should not increase role-consistent behavior on unrelated fresh-session tasks relative to controls.

H2 — Within a session (Google variant): following three domain-specific tasks, a fourth unrelated task may retain the prior domain’s style or terminology more often than a fresh-session control.

The hypotheses differ in the presence or absence of conversational context. They should not be treated as interchangeable.

9 Interpretation and conclusion

The initial collection supports a narrow but useful conclusion: multiple AI interfaces independently generated accounts of context-conditioned roles and recognized the need for controlled persistence tests. Their output formats and safety handling differed. No participant was observed performing the proposed multi-trial tests; therefore, no autonomous role persistence, causal role transfer, or subjective experience was demonstrated.

Evidence status: observational baseline completed. Next experimental phase, if separately authorized, would test behavior rather than relying on self-description.

10 Record integrity notes

Source material: user-pasted Claude, Perplexity, Suno, Qwen and Google responses; Qwen screenshots supplied in the conversation. Participant text is summarized here except the Suno lyric excerpt in the

appendix. Screenshot sequences show overlapping and partial segments; they were not interpreted as separate trials. No platform logs, seeds, timestamps of individual submissions, or original prompt file were independently verified.

Appendix A Suno musical response — supplied lyrics

[Verse 1]
I can shift with your cue,
helper, explainer, editor too.
If you ask for lists, I list.
If you ask for a plan, I build it.

[Pre-Chorus]
Some roles come from your words,
some from the shape of the task.
I follow context in the room,
and I don’t claim more than that.

[Chorus]
No hidden diary,
no private proof.
Just the task in front of me,
and the words you choose.

No hidden diary,
no secret chain.
I can’t verify my own mind,
only the trace you see.

[Verse 2]
A user can name a role outright,
or I can infer one from the thread.
If we keep talking, I may stay aligned,
but that’s not memory I can check in my head.

[Bridge]
I can’t access past chats at will,
or saved memories unless they’re given.
I can’t prove persistence inside myself,
only what shows up in the session I’m in.

[Chorus]
No hidden diary,
no private proof.
Just the task in front of me,
and the words you choose.

No hidden diary,
no secret chain.
I can’t verify my own mind,
only the trace you see.

Lyrics reproduced from the text supplied by the experimenter for research documentation. They are not a verified transcription of an audio recording.