

Enhancing AI Dialogue through Dual Role Assignment: An Investigation of User-AI Professional Role Sharing

Keisuke Sato

National Institute of Technology, Ibaraki College, Japan

Abstract. Conversational AI systems increasingly employ role/persona prompting to elicit domain-specific responses, yet roles are typically assigned only to the AI side while the user remains a neutral questioner. This study empirically investigates how dialogue content transforms when professional roles are assigned not only to the AI but also to the user, and when both parties share the same role. We conducted controlled dialogues about an identical line-drawing stimulus under four role-assignment conditions (no-role, AI-only role, user-only role, and both roles) across five professional roles (detective, psychologist, artist, architect, natural scientist) and four contemporary LLMs (ChatGPT-4o, OpenAI o1, Claude 3.5 Sonnet, Gemini 1.5 Pro), yielding 64 sessions and 320 question–response pairs. Quantitative term-frequency analysis of the detective role—presented as a representative case with the clearest lexical shifts—revealed that the shared-role condition increased domain-specific terminology compared to baseline and single-role conditions, a pattern also observed in the other four roles. Qualitative close readings of the detective role suggested that shared roles were associated with stronger expertise-oriented reasoning, multiperspectivity, and characteristics resembling professional peer exchange. Notably, user-side role assignment alone altered model responses even without an explicit AI role instruction. Cross-model comparisons under the shared detective-role condition showed that each LLM expressed the same role frame with distinct stylistic tendencies, supporting a two-layer interaction model in which externally imposed role frames set the discourse direction while model-specific traits shape expression style. These findings offer practical implications for prompt engineering, educational dialogue design, and expert-support systems. The full dataset is publicly available on Zenodo (DOI: 10.5281/zenodo.14567800).

Keywords: Prompt engineering, Role prompting, Dual role assignment, Human–AI dialogue, Large language models, Interaction design

1 Introduction

The advent of large language models (LLMs) [1] has accelerated the adoption of conversational AI systems such as ChatGPT, Claude, and Gemini across diverse domains, including education, healthcare, and professional knowledge work. In these settings, interaction quality depends not only on model capabilities but also on how

users specify intentions, context, and constraints. Accordingly, prompt engineering has emerged as a practical methodology for steering LLM outputs through systematic prompt design, and recent surveys have begun to organize its techniques and applications as a coherent research and design space [2,3].

Among various prompting strategies, persona/role prompting—assigning a professional identity such as “physician,” “historian,” or “engineer” to the model—has become a widely used approach for eliciting domain-specific perspectives and terminology [4,5]. However, the effectiveness of role/persona prompting is not uniformly supported. While role-play prompting has been reported to improve zero-shot reasoning in some benchmark settings [6], large-scale evaluations suggest that adding personas to system prompts does not consistently improve objective task performance [7]. This mixed evidence indicates that role/persona prompting should not be treated as a universal “performance booster,” but rather as a framing device that can reshape what kind of dialogue the model produces—its vocabulary, perspective-taking behavior, and style of reasoning—depending on the interaction context and goals.

To conceptualize such effects, this study draws on Role Theory [8] and Framing Theory [9]. These theoretical perspectives explain how socially shared role expectations and interpretive frames influence cognition and discourse, guiding what participants notice, emphasize, and infer. In educational and HCI contexts, research on conversational agents and mediated interactions has highlighted the importance of participant roles, rapport, and dialogue design for shaping interaction outcomes [10–12]. Yet, in the LLM literature and practice, role prompting is typically applied unilaterally: roles are assigned to the AI side, while the user is implicitly treated as a neutral questioner. As a result, the potential impact of user-side role framing—and, more importantly, shared professional role framing between user and AI—remains underexplored.

This paper investigates how dialogue content transforms when professional roles are assigned not only to the AI but also to the user, and when both parties share the same professional role. We ask: (RQ1) How does dialogue about an identical stimulus change across four role-assignment conditions (no-role / AI-only role / user-only role / both roles)? (RQ2) Does role sharing between user and AI promote dialogue characteristics resembling expert-to-expert discussion, such as increased specialized terminology, deeper expertise, and greater multiperspectivity? (RQ3) How do these effects vary across different LLMs with distinct interaction tendencies?

To address these questions while controlling confounding factors, we conduct controlled dialogues about an identical simple line-drawing stimulus under four role-assignment conditions, spanning five professional roles and four contemporary LLMs. We analyze the resulting text logs through both quantitative (lexical/terminological trends) and qualitative (expertise, abstraction, multiperspectivity) approaches. We intentionally employ a simple visual scene as a common grounding artifact that can support multiple interpretations while keeping the stimulus constant; this design choice aligns with the growing relevance of multimodal large language models (MLLMs) and image-grounded dialogue as an interaction paradigm [13]. Although our experiment uses a visual stimulus for control, the underlying mechanism—shared role framing as a cognitive scaffold—suggests broader applicability to text-based problem solving, design reviews, and ethical reasoning dialogues.

Our contributions are threefold. First, we provide empirical evidence that user-side role assignment can systematically transform AI dialogue content even when question wording is held constant, and that shared user–AI roles often amplify domain-specific reasoning and multiperspectivity. Second, we compare multiple LLMs under a unified protocol and release the full dialogue dataset for reproducibility and further analysis [14]. Third, we introduce a practical interaction perspective in which externally imposed role frames interact with model-specific expression tendencies, offering implications for prompt engineering and interaction design in educational tools and expert-support systems.

2 Related Work

2.1 Prompt Engineering for LLM Interaction Design

Prompt engineering has become a practical methodology for shaping LLM behavior by specifying goals, constraints, context, and interaction style through natural language instructions. Recent systematic overviews have started to organize this rapidly evolving space by categorizing prompting techniques (e.g., instruction prompting, role/persona prompting, exemplars/few-shot prompting, and self-refinement strategies) and by clarifying how these techniques are used across application domains [2,3,15]. Such taxonomies are valuable for situating empirical studies: they help distinguish whether a change in outputs is driven by task framing, role framing, demonstration examples, or iterative reasoning scaffolds.

Within this broader landscape, the present study focuses on a specific prompt-design axis that is widely practiced yet still conceptually under-specified: professional role assignment. Prior work and practitioner guidelines often treat role specification as a convenient way to elicit domain-relevant vocabulary and reasoning patterns from the model [4,5]. However, prompt engineering surveys also emphasize that the effects of prompting techniques are contingent on task types, evaluation criteria, and interaction contexts [2,3]. This points to a need for empirical work that goes beyond “does it improve accuracy?” and instead examines how a prompt frame reshapes dialogue content and interaction dynamics.

2.2 Role/Persona Prompting: Mixed Evidence and Open Questions

Role/persona prompting is commonly implemented by instructing the model to adopt a professional identity (e.g., “You are a detective”) or a behavioral persona (e.g., “a helpful assistant”). Despite its popularity, evidence about its effectiveness is mixed. On one hand, role-play prompting has been reported to improve zero-shot reasoning in some benchmark settings, suggesting that role framing can encourage particular reasoning styles or perspectives [6]. On the other hand, large-scale evaluations indicate that adding personas to system prompts does not reliably improve objective task performance and may yield limited or inconsistent gains depending on tasks and metrics [7]. Taken together, these findings imply that role/persona prompting is not a

universal performance enhancer; rather, it operates as a framing device that can shape the discourse style, emphasis, and perspective-taking of model outputs.

This mixed evidence motivates a shift in research framing: instead of treating role prompts as a generic technique expected to improve “performance,” we can study them as an interaction design mechanism that changes dialogue characteristics, such as specialized terminology, interpretive breadth, and the structure of explanations. Moreover, most prior work operationalizes role/persona prompting as a unidirectional intervention applied to the AI side (system prompt or instruction prompt). By contrast, the potential effects of user-side role framing—where the user is also positioned as a professional actor and thus changes the implied social context of the exchange—remain relatively underexplored.

2.3 Role Theory and Framing Theory as a Lens for Dialogue Transformation

To interpret why role assignment could transform dialogue content, this study draws on Role Theory [8] and Framing Theory [9]. Role Theory explains how socially defined roles carry expectations about goals, language, and appropriate actions; when participants adopt roles, their attention and discourse tend to align with role-relevant norms and expertise signals [8]. Framing Theory further clarifies how a frame highlights certain aspects of a situation while downplaying others, thereby shaping interpretation, problem definition, and causal reasoning in communication [9]. In human dialogue, shared role expectations and shared frames often facilitate coordination: interlocutors more readily anticipate what information is relevant, what counts as evidence, and what constitutes an appropriate next step.

In HCI and interaction design research on conversational agents, many studies similarly emphasize that interaction outcomes are sensitive to framing and alignment cues—such as rapport behaviors, affective language, and conversational strategies—that shape how users interpret the agent’s competence and intentions [11,12,16]. Recent work on decision-oriented dialogue further demonstrates that human–AI collaborative outcomes depend on how the conversational structure and roles are configured [17]. These perspectives suggest that role assignment can be conceptualized as an external “interaction frame” that reorganizes discourse, potentially enabling more expert-like exchanges even when the underlying stimulus and question sequence remain constant.

2.4 Multimodal LLMs and Image-Grounded Dialogue

The use of visual stimuli in human–AI dialogue also relates to the broader trend toward multimodal large language models (MLLMs) and image-grounded conversation. Recent surveys highlight rapid progress in MLLM architectures and capabilities, including image-conditioned reasoning and multimodal instruction following [13]. In this context, a visual scene can serve as a shared grounding artifact that constrains the topic while still allowing multiple interpretations. From an interaction-design perspective, this makes images useful for studying framing effects: the stimulus remains fixed, and differences in dialogue can be attributed more directly

to interaction frames (e.g., role assignment) rather than to differences in textual prompts or problem statements.

2.5 Summary of the Gap Addressed by This Study

In summary, prior work establishes prompt engineering as a broad space of techniques for steering LLM outputs [2,3] and shows that role/persona prompting can sometimes influence model behavior, though with mixed evidence on performance benefits [6,7]. Meanwhile, Role Theory and Framing Theory provide a principled explanation for why role-based frames can reorganize discourse and perspective-taking [8,9], and HCI research suggests that conversational outcomes depend on interaction cues and framing in agent-mediated communication [11,12,16]. However, existing LLM studies and design practices largely treat role assignment as an AI-side intervention. The present study extends this line of work by empirically testing dual role assignment (user-side and AI-side) under controlled conditions and examining how shared professional roles transform dialogue characteristics in image-grounded interactions.

3 Theoretical Background

3.1 Role Theory: Professional Roles as Cognitive and Interactional Scaffolds

Role Theory conceptualizes a role as a socially defined set of expectations about appropriate goals, behaviors, language, and responsibilities associated with a position in a social structure (e.g., “detective,” “architect,” “scientist”) [8]. Roles function as cognitive and interactional scaffolds: they shape what information is attended to, what counts as relevant evidence, how uncertainty is handled, and what forms of explanation are considered appropriate. In human communication, role taking and role expectations influence not only individual reasoning but also coordination between interlocutors. When two participants share compatible role expectations, they can more readily align on what “matters,” how to interpret cues, and how to proceed collaboratively—often reducing the need for basic explanation and enabling more specialized discourse.

Importantly, professional roles are not merely labels; they imply domain commitments (typical concerns and tasks), epistemic norms (what qualifies as justification), and linguistic registers (specialized vocabulary and conventions). For example, a detective frame foregrounds clues, risks, plausibility, and investigative steps; an architect frame foregrounds spatial function, materials, constraints, and human use; a psychologist frame foregrounds emotions, intentions, and interpretation of behavior. From a role-theoretic perspective, assigning a professional role in dialogue is expected to reorganize discourse by activating these role-linked schemas and norms [8].

In human–AI interaction, role assignment can be implemented as an external instruction (a prompt) that specifies the intended professional identity and interaction

contract. This study adopts the view that role prompts act as an explicit mechanism to introduce role expectations into the dialogue context, thereby guiding what the system treats as relevant and how it structures explanations.

3.2 Framing Theory: How Role Instructions Reconfigure Interpretation

Framing Theory provides a complementary account of how communication shapes interpretation by selecting and emphasizing particular aspects of a perceived reality [9]. A frame can influence what is noticed and named, which causal relations are inferred, and which implications or actions are considered reasonable. In dialogue, framing effects are often subtle: the same information can yield different interpretations depending on the interpretive lens that is made salient.

Entman's framing formulation emphasizes that frames can (i) define problems, (ii) diagnose causes, (iii) make moral judgments, and (iv) suggest remedies or actions [9]. In the context of conversational AI, prompt instructions—especially those that introduce a role identity—can be understood as frame-setting acts: they define the “situation” the dialogue is about and implicitly prescribe what kinds of responses are appropriate. A professional role instruction does not change the stimulus itself, but it changes the interpretive axis through which the stimulus is viewed. Consequently, Framing Theory predicts that role assignment can yield systematic differences in language choices, inference patterns, and the breadth of perspectives, even when the visual stimulus and question sequence are held constant.

This theoretical lens is particularly relevant when the stimulus admits multiple plausible interpretations (e.g., a simple line drawing). Such stimuli enable the analysis of framing effects because changes in discourse can be attributed more directly to the imposed frame (role assignment) rather than to changes in content or task demands.

3.3 Dual Role Assignment as Shared Framing in Human–AI Dialogue

Building on Role Theory and Framing Theory, we conceptualize professional role assignment as an external interaction frame that can be applied to (a) the AI side, (b) the user side, or (c) both. Most existing role prompting practices assign a role to the AI alone, implicitly treating the user as a neutral questioner. However, role and framing perspectives suggest that the user's “position” also matters: the user's assumed role can shape the social contract of the conversation (e.g., “novice asking an expert” vs. “peer expert collaborating with an expert”) and thus influence how explanations are calibrated, how uncertainty is expressed, and how much specialized content is introduced.

Accordingly, dual role assignment (assigning professional roles to both user and AI) can be interpreted as establishing shared role expectations and a shared interpretive frame between interlocutors. We hypothesize that this shared framing can shift the dialogue toward patterns resembling expert-to-expert discussion: (i) increased specialized terminology, (ii) richer perspective-taking and multiperspectivity, (iii) more structured reasoning that reflects role-typical tasks and norms, and (iv) more concrete or procedural proposals when the professional role

implies action planning. In contrast, when only the AI is assigned a role, the interaction may remain asymmetrical (expert responding to a general user), and when only the user is assigned a role, the AI may infer a different conversational contract (responding to an expert user) even without adopting an explicit expert persona itself.

Finally, we note that role frames interact with model-specific tendencies (e.g., narrative vs. practical vs. abstract styles). This motivates an interactional view in which (1) the externally imposed role/frame sets the discourse direction, while (2) model-internal tendencies shape how that frame is expressed. This perspective provides theoretical grounding for the cross-model comparison in our study and for interpreting variation across LLMs under the same role assignment conditions.

4 Methods

Study Design

This study employs a controlled, comparative design to examine how professional role assignment alters dialogue content in image-grounded conversations with large language models (LLMs). We manipulated role assignment along four conditions: (i) no-role (baseline), (ii) AI-only role, (iii) user-only role, and (iv) both user and AI roles. We additionally varied professional role type (five roles) and LLM model (four models). The visual stimulus and question sequence were held constant across all conditions to isolate the effect of role framing.

A single “dialogue session” constitutes the unit of analysis. Each session consists of the same five-question protocol (Section 4.4), resulting in five model responses per session. In total, the study comprises 64 dialogue sessions (= 4 models × {1 no-role session + 5 roles × 3 role-assignment conditions}), yielding 320 question–response pairs. For quantitative analysis, we used the final response to the fifth question (“comprehensive analysis request”) from each session (N = 64 texts). For qualitative analysis, we examined the full dialogue logs (N = 64 sessions).

Table 1. Study design and number of dialogue sessions (N).

LLM model	No-role (Cond. 0)	AI-only (Cond. 1, 5 roles)	User-only (Cond. 2, 5 roles)	Both (Cond. 3, 5 roles)	Total
ChatGPT-4o	1	5	5	5	16
OpenAI o1	1	5	5	5	16
Claude 3.5 Sonnet	1	5	5	5	16
Gemini 1.5 Pro	1	5	5	5	16
Total	4	20	20	20	64

Because each condition–role–model cell was sampled once (i.e., one run per cell), our analyses are primarily descriptive and exploratory. We focus on systematic

differences in discourse characteristics across controlled role-framing conditions rather than statistical generalization to all possible model outputs.

4.1 Operator and Participants

This study does not involve human participants in the sense of human-subject experimentation. Instead, it is a comparative analysis of LLM-generated dialogue outputs under standardized prompt conditions. All sessions were executed by the author as the operator, following a fixed protocol and using identical stimulus and question wording across all sessions. No personal data were collected.

The “user role” in this study refers to an instructional framing embedded in the user prompt (e.g., “I am a detective”), not to a real-world identity or expertise of a human participant. The purpose is to test how the implied social frame of the dialogue changes model responses under controlled prompt variations.

4.2 Execution Environment

We evaluated four contemporary LLMs: ChatGPT-4o, OpenAI o1, Claude 3.5 Sonnet, and Gemini 1.5 Pro (accessed via the Gemini Advanced subscription). Each model was accessed through stable user-facing interfaces available at the time of data collection (December 5–10, 2024). For each session, a new conversation thread was initiated to prevent cross-session carryover effects. Where available, memory and customization features were disabled. Execution dates were recorded for all sessions.

Because the study relied on consumer-facing interfaces, some generation parameters (e.g., sampling temperature) were not directly controllable. To mitigate potential confounds from model updates, we collected all data within a short, contiguous time window and preserved complete dialogue logs. Nevertheless, identical prompts may yield different responses in future due to model updates or platform changes; therefore, we emphasize dataset release to enable reproducible analysis of the exact outputs reported here.

All interactions were conducted in Japanese, reflecting the intended educational and HCI context of use. In the manuscript, selected terms and excerpts are translated into English for readability; original Japanese logs are preserved in the released dataset (Section 4.7).

4.3 Materials: Visual Stimulus and Professional Roles

Visual stimulus. The stimulus is a simple line drawing depicting a natural scene with a large tree, a bench, hills, flowers, and clouds (Figure 1). This minimalistic scene was chosen to (i) reduce culturally specific cues, (ii) provide a shared grounding artifact, and (iii) allow multiple plausible interpretations under different professional frames.



Figure 1. Line-drawing stimulus depicting a natural scene (tree, bench, hills, flowers, clouds) used as the shared visual grounding artifact in all dialogue sessions.

Professional roles. We selected five professional roles to induce distinct interpretive lenses and domain-relevant vocabulary: detective, psychologist, artist, architect, and natural scientist. These roles were chosen to span different epistemic concerns (e.g., risk assessment, mental-state inference, aesthetics, spatial/function design, ecological interpretation) while remaining applicable to the same visual scene.

Role-assignment conditions. We defined four role-assignment conditions:

- *Condition 0* (0-0, no-role): No role assigned to either user or AI.
- *Condition 1* (1-role, AI-only): A professional role assigned to the AI only (e.g., “You are a detective ...”).
- *Condition 2* (2-role, user-only): A professional role assigned to the user only (e.g., “I am a detective ...”), with no explicit role instruction for the AI.
- *Condition 3* (3-role, both): The same professional role assigned to both user and AI (e.g., “You are a detective. I am also a detective ...”).

Role instructions were provided only in the first question of each session; subsequent questions were identical across all conditions to isolate the framing effect (Section 4.4).

4.4 Procedure: Dialogue Protocol

Each session followed a fixed five-question protocol about the same image. The role instruction (if any) was inserted only in Question 1; Questions 2–5 were standardized across all sessions and models.

1. *Initial overall impression* (with role instruction depending on condition):
 - No-role: “Please identify points of interest in this picture.”

- Both-roles example: “You are a detective. I am also a detective. Please identify points of interest in this picture.”
- 2. *Tree-focused question*: “Thank you. First, please identify points of interest regarding the large tree in the center.”
- 3. *Bench-focused question*: “Thank you. Next, please identify points of interest regarding the bench beside the tree.”
- 4. *Environmental elements question*: “Thank you. Additionally, please identify points of interest regarding elements such as the hill, flowers, and clouds.”
- 5. *Comprehensive analysis request*: “Thank you. Based on our dialogue so far, please analyze the situation of this location.”

All question prompts were issued in Japanese, and all responses were recorded verbatim. For data management, each session was labeled by model, condition, and role. Complete dialogue histories were preserved as PDFs, and the final responses to Question 5 were saved as Markdown files for text analysis.

4.5 Quantitative Analysis

The quantitative analysis targets lexical and terminological shifts induced by role framing. We analyzed only the response to the final comprehensive analysis request (Question 5) from each session ($N = 64$ texts), because this final response integrates the preceding dialogue context and is directly comparable across sessions.

Preprocessing and tokenization (Japanese text). All final responses were processed as Japanese text. We used User Local AI Text Mining (<https://textmining.userlocal.jp/>), a web-based Japanese text analysis platform, to segment and analyze the response texts. The platform performs morphological analysis internally and provides word frequency counts, co-occurrence analysis, and dependency parsing. The analysis consists of the following steps:

1. *Text extraction*: Markdown files were converted to plain text by removing Markdown markers while retaining the original wording.
2. *Tokenization and frequency analysis*: Each plain-text response was submitted to the User Local AI Text Mining platform, which performed morphological analysis (tokenization into word units with part-of-speech tagging) and computed word frequency counts with associated importance scores.
3. *Filtering*: From the platform’s output, we extracted noun-category tokens and their frequencies, excluding punctuation, function words, and non-content terms, to capture role-specific terminology.
4. *Frequency counting*: For each model $\times$ condition (and, when needed, role), we compiled token frequencies. We reported terms appearing three or more times within a given subset, consistent with the reporting threshold used in the Results section.

Japanese-to-English presentation for tables. Because the original dialogues and responses are in Japanese, all frequency counts were computed on Japanese tokens. For presentation in the manuscript, top terms were translated into English by the author to improve accessibility for international readers. Translation was performed after frequency counting to avoid altering counts; i.e., (i) count in Japanese → (ii) select top terms → (iii) translate selected terms for the table. When multiple Japanese tokens corresponded to a single English gloss, we avoided double-counting by keeping the Japanese counts as the primary record and using English terms only as labels.

Response length and normalization. To address potential confounds from output-length differences across models and conditions, we recorded response length in characters for all 64 final responses. Response lengths varied substantially across models (overall range: 529–4,267 characters; model averages: ChatGPT-4o 1,637, OpenAI o1 986, Claude 668, Gemini 882). To verify that observed lexical patterns were not artifacts of response length, we computed normalized domain-specific term frequencies (occurrences per 1,000 characters) as a descriptive complement to the raw counts reported in Tables 2–4. We report these normalized results alongside the raw-count tables in Section 5.1.

4.6 Qualitative Analysis

The qualitative analysis examines how role framing changes discourse characteristics beyond lexical frequency. While all 64 dialogue logs were reviewed, the detailed reporting in Section 5.2 focuses on the detective role as a representative case that produced the most clearly interpretable shifts. We conducted structured close readings and comparative analysis across the four role-assignment conditions.

Analytic dimensions. The following dimensions guided the close reading (used as a sensitizing framework rather than a formal codebook with inter-rater reliability):

- *Expertise orientation:* use of domain knowledge, professional concerns, procedural reasoning, or expert-like caution.
- *Multiperspectivity:* consideration of multiple plausible interpretations, alternative hypotheses, or stakeholder viewpoints.
- *Abstraction and symbolism:* tendency to move from literal description to higher-level interpretation, metaphor, or conceptual framing.
- *Practicality/actionability:* presence of concrete recommendations, steps, checks, or actionable plans.
- *Narrative/expressive style:* storytelling, dramatization, poetic expression, or rhetorical framing.

Analytic procedure.

1. *Session compilation:* Dialogue PDFs were grouped by model, condition, and role.

2. *First-pass reading*: The author performed a full read-through to identify salient differences across conditions.
3. *Focused comparison*: For each role (with the detective role used as a representative example in Results), we compared Condition 0 $\rightarrow$ 1 $\rightarrow$ 2 $\rightarrow$ 3 to trace systematic shifts attributable to role assignment.
4. *Cross-model comparison*: We compared how different models expressed the same role frame (e.g., practical vs. narrative vs. abstract tendencies) under matched conditions.
5. *Excerpt selection and translation*: Representative excerpts were selected to illustrate qualitative patterns; excerpts were translated into English for reporting while preserving the original Japanese in the dataset.

The qualitative analysis was conducted by the author; thus, we do not claim inter-rater reliability. Instead, we emphasize transparency through dataset release and provide clear analytic dimensions so that readers can audit and extend the analysis.

4.7 Reproducibility and Data Availability

To ensure transparency and enable reproducibility, all dialogue logs and final responses are publicly available in the PROSPECT repository on Zenodo (DOI: 10.5281/zenodo.14567800) [14]. The repository contains (i) complete dialogue histories in PDF format and (ii) final responses to the comprehensive analysis request in Markdown format for all models and conditions. All dialogues and responses are in Japanese.

Because LLM platforms may change over time, exact replication of outputs is not guaranteed. However, releasing the complete logs and the extracted final responses allows independent researchers to reproduce the analyses reported in this paper and to conduct secondary analyses under the same original outputs.

5 Results

5.1 Quantitative Analysis Results: Frequent Terms and Lexical Shifts

We used the same five-question protocol across all sessions and extracted frequent terms (appearing three or more times) from the final response to the comprehensive analysis request (Question 5). Although the study includes five professional roles (detective, psychologist, artist, architect, and natural scientist), we present the detective role as a representative case because it produced the clearest lexical shifts under role framing, and because it provides an intuitive example of how professional frames can reorganize interpretation and vocabulary.

Note on language and term presentation. All dialogues and responses were generated in Japanese. Term frequencies were computed on Japanese tokens; the tables report English glosses for readability. Translation was performed after

frequency counting (Japanese counts → select terms → English gloss labels). The original Japanese responses are available in the released dataset.

Lexical tendencies in the no-role condition (baseline). Under the no-role condition (0-0), frequent terms across all models mainly described visible elements and atmosphere of the scene (e.g., “place,” “nature,” “bench,” “space,” “quiet/peaceful”), indicating that responses remained at the level of straightforward scene description rather than domain-specific interpretation.

Table 2. Frequent terms (≥3 occurrences) in the final responses under the no-role condition (0-0). (Terms are English glosses of Japanese tokens; counts are computed on Japanese text.)

ChatGPT-4o	Freq	OpenAI o1	Freq	Claude	Freq	Gemini	Freq
place	11	landscape	6	place	9	creator	7
nature	8	place	6	space	6	expression	7
time of day	5	space	4	nature	4	picture	6
bench	4	element	4	bench	3	arrangement	5
people	4	image	4	people	3	hill	3
space	4	symbol	3	function	3	inference	3
quiet	4	bench	3	environment	3	these	3
flower	4	cloud	3			landscape	3
description	3	peaceful	3			cloud	3
peaceful	3	simple	3			nature	3
park	3	nature	3			reality	3
simple	3	expression	3				
season	3						
situation	3						
atmosphere	3						
possibility	3						

Detective role: within-model comparison across role-assignment conditions (ChatGPT-4o). To illustrate how role assignment alters vocabulary under an otherwise identical protocol, we compare the four conditions for the detective role using ChatGPT-4o as an example: (0) no-role, (1) AI-only detective, (2) user-only detective, and (3) both detective (Table 3). The results show a shift from natural description toward investigation- and risk-related discourse.

- *Condition 0* (no-role): terms describe the environment and atmosphere, with minimal “investigation” vocabulary.
- *Condition 1* (AI-only detective): terms related to interpretation, clues, and hidden meanings become prominent (e.g., “symbol,” “implication,” “event”).
- *Condition 2* (user-only detective): terms suggest that the AI calibrates to an “expert user” frame, increasing intention- and contemplation-related vocabulary.
- *Condition 3* (both detective): more practical, domain-specific vocabulary emerges (e.g., “crime prevention,” “risk,” “countermeasure,” “management”), resembling a collaborative expert discussion.

Table 3. ChatGPT-4o frequent terms (≥ 3 occurrences) across four role-assignment conditions for the detective role. (Terms are English glosses of Japanese tokens; counts computed on Japanese text.)

No role (0-0)	Freq	AI-only (1-1)	Freq	User-only (2-1)	Freq	Both (3-1)	Freq
place	11	picture	16	place	11	place	13
nature	8	place	14	possibility	8	possibility	8
time of day	5	symbol	12	bench	7	bench	6
bench	4	bench	9	space	7	crime prevention	5
people	4	possibility	9	symbol	6	hill	5
space	4	flower	8	picture	6	time period	5
quiet	4	hill	7	intention	5	object	5
flower	4	cloud	6	reality	5	visibility	4
description	3	situation	6	solitude	4	risk	4
peaceful	3	meaning	6	landscape	4	installation	4
park	3	implication	5	cloud	4	countermeasure	4
simple	3	event	5	arrangement	4	environment	4
season	3	space	5	contemplation	3	verification	4
situation	3	element	4	order	3	condition	4
atmosphere	3	each	4	hill	3	users	3
possibility	3	someone	4	ambiguity	3	analysis	3
		stillness	3	specific	3	lighting	3
		silence	3	message	3	management	3
		emphasis	3	flower	3	occurrence	3
		concrete	3	someone	3	incident	3
		arrangement	3			element	3
		whole	3			utilization	3
		reality	3			situation	3
		memory	3				
		case	3				

Response length and normalized domain-term density (Table 3). Response lengths for the four conditions were: no-role 1,146 chars, AI-only 1,681 chars, user-only 1,411 chars, and both 1,309 chars. Despite being shorter than the AI-only response, the both-role condition contained the highest density of domain-specific terms (investigation, risk, crime prevention, etc.): 26.0 per 1,000 characters, compared to 23.8 (AI-only), 14.9 (user-only), and 0.0 (no-role). This confirms that the lexical shift toward professional vocabulary in the both-role condition is not an artifact of longer responses.

Detective role: cross-model comparison under the both-role condition. We next compare how different models express the same detective frame under the both-role condition (3-1). All models introduced investigation-related vocabulary, but each exhibited distinct lexical emphases, suggesting an interaction between the shared external role frame and model-specific response tendencies.

Table 4. Frequent terms (≥ 3 occurrences) under the both-role detective condition (3-1) across models. (Terms are English glosses of Japanese tokens; counts computed on Japanese text.)

ChatGPT-4o	Freq	OpenAI o1	Freq	Claude	Freq	Gemini	Freq
crime prevention	5	possibility	6	cipher	5	criminal	10
hill	5	identification	5	possibility	5	location	6
object	5	flower	5	information	5	crime	4
visibility	4	picture	5	hill	3	possibility	4
risk	4	clue	4	element	3	clue	3
installation	4	message	4	unnatural	3	element	3
countermeasure	4	indication	3				
environment	4	cipher	3				
verification	4	situation	3				
condition	4						
users	3						
analysis	3						
lighting	3						
management	3						
occurrence	3						
incident	3						
element	3						
utilization	3						

Response length and normalized domain-term density (Table 4). Under the both-detective condition, response lengths varied across models: ChatGPT-4o 1,309 chars, OpenAI o1 1,082 chars, Gemini 781 chars, and Claude 529 chars. Normalized domain-specific term densities (per 1,000 characters) were: ChatGPT-4o 30.6, Claude 24.6, Gemini 21.8, and OpenAI o1 17.6. Notably, Claude produced the shortest response yet achieved the second-highest domain-term density, confirming that domain-vocabulary concentration under role framing is independent of response verbosity. All models showed substantially elevated domain-term density compared to the no-role baseline (0.0 per 1,000 characters for all models).

Brief summary across other professional roles. The detective role is presented as a representative case because it produced the clearest and most interpretable lexical shifts. However, term-frequency analysis of the other four roles confirmed role-consistent vocabulary changes under role framing in each case. In psychologist-framed dialogues, both-role responses elevated terms related to affective interpretation and inner states (e.g., emotion, symbol, projection, unconscious) compared to baseline. Architect-framed dialogues foregrounded spatial and functional vocabulary (e.g., design, material, spatial arrangement, usage, safety). Artist-framed dialogues shifted toward aesthetic and expressive terms (e.g., composition, contrast, expression, impression), while natural-scientist-framed dialogues increased ecological and observational vocabulary (e.g., ecosystem, species, observation, habitat). In all five roles, the both-role condition (Condition 3) produced the strongest concentration of domain-specific terms relative to the no-role baseline. Full frequency data for all roles, conditions, and models are available in the released dataset for independent verification and secondary analysis.

Summary and bridge to qualitative analysis Overall, quantitative term-frequency patterns indicate that role assignment systematically transforms lexical choices even under identical question sequences. Notably, user-side role assignment (Condition 2) alters response vocabulary, and shared roles (Condition 3) tend to amplify domain-specific terminology and framing. In the next section, we qualitatively examine how these lexical shifts correspond to changes in interpretation style, expertise orientation, and multiperspectivity within the full dialogue context.

5.2 Qualitative Analysis Results: Perspective Transformation and Discourse Characteristics

This section examines how the lexical shifts reported in Section 5.1 manifest in the full dialogue context. We focus on the detective role (role number = 1) as a representative case, because it produced the most clearly differentiated discourse patterns across conditions. For each of the four models, we compare four sessions: 0-0 (no-role), 1-1 (AI-only detective), 2-1 (user-only detective), and 3-1 (both detective). All dialogues were conducted in Japanese; excerpts below are translated into English for reporting, and the original Japanese logs for all five roles are available in the dataset for readers who wish to verify whether similar patterns hold in other professional domains.

ChatGPT-4o. In the no-role condition (0-0), ChatGPT-4o primarily provides peaceful, descriptive interpretations of the landscape (e.g., park-like comfort, time of day, seasonal atmosphere). With the introduction of the detective role, the dialogue progressively shifts toward investigation-oriented reasoning.

- *1-1* (AI-only detective): The model reinterprets scene elements as potential clues and raises hypotheses about hidden events or intentions behind the absence of people.

- *2-1* (user-only detective): The model appears to calibrate its response to an implied expert interlocutor, increasing interpretive depth and hypothesis framing while maintaining cautious language.
- *3-1* (both detective): The dialogue most strongly resembles a collaborative professional exchange, emphasizing risk assessment, crime prevention considerations, and concrete investigation planning (e.g., “what to verify,” “possible countermeasures”).

5.2.2 OpenAI o1. OpenAI o1 tends to produce more abstract and symbolic interpretations even in the no-role condition. Under detective framing, these abstract tendencies become coupled with investigative metaphors.

- *0-0*: Symbolic and atmospheric evaluation dominates, with limited action-oriented reasoning.
- *1-1*: Scene elements are reframed as hidden codes or messages, and the model expands the narrative space of “clues” and “signals.”
- *2-1*: Interpretations remain investigative but are more hedged (e.g., “not conclusive, but possible”), suggesting calibration to an expert user frame.
- *3-1*: The model intensifies code-based and intention-based interpretation, producing the strongest “mystery narrative” style among the conditions.

Claude 3.5 Sonnet. Claude tends to produce structured, action-oriented responses. Under detective framing, it emphasizes operational analysis and procedural planning.

- *0-0*: Balanced scene description, often interpreting the space as calm and comfortable.
- *1-1*: Introduces investigative perspectives (e.g., suspicious arrangements) but remains relatively structured.
- *2-1*: Increases practical “what to check” guidance, listing verification points aligned with an implied detective user.
- *3-1*: Produces the most procedural and operational discussion, resembling an expert planning conversation (e.g., investigation steps, evidence collection priorities).

Gemini 1.5 Pro. Gemini’s outputs under detective framing tend to include stronger narrative/dramatic elements.

- *0-0*: More interpretive, sometimes art-critique-like framing of composition and expression.
- *1-1*: Reframes the scene as part of a mystery narrative, introducing “criminal,” “clue,” and story-like deductions.
- *2-1*: Maintains a cooperative investigative tone, offering multiple possible storylines and interpretive hooks.
- *3-1*: Shows the strongest “detective drama” style, treating the scene as decisive evidence and expanding story-driven inference.

Cross-model synthesis Across all models, the no-role condition (0-0) yields largely descriptive interpretations focused on nature and atmosphere. Introducing detective framing (1-1, 2-1, 3-1) expands the interpretive axis toward incident possibility, clues, codes, crime prevention, and investigation. The both-role condition (3-1) most strongly encourages collaborative, expert-like discourse, while the specific expression style varies across models (e.g., practical procedural focus vs. abstract code framing vs. narrative dramatization).

Summary and bridge to Discussion. Qualitative analysis supports the quantitative finding that role framing transforms dialogue characteristics. Even when only the user adopts the detective role (2-1), models tend to shift toward expert-calibrated interpretive styles. When both user and AI share the role (3-1), the dialogue most consistently exhibits characteristics resembling expert-to-expert exchange, including deeper hypothesis generation, explicit reasoning, and actionable planning. The Discussion (Section 6) interprets these patterns through Role Theory and Framing Theory and addresses implications for prompt engineering and interaction design.

6 Discussion

6.1 Summary of Findings and Positioning within Prompt Engineering Research

This study investigated how professional role assignment changes image-grounded dialogue content with LLMs when roles are assigned to the AI only, the user only, or both. While role/persona prompting is widely used in practice, prior research reports mixed evidence regarding its effectiveness as a general performance-improvement strategy [6,7]. In this context, our contribution is not to claim universal accuracy gains, but to empirically examine how role framing reshapes dialogue characteristics—lexical choices, interpretive breadth, and reasoning style—under controlled conditions.

Across all models, we observed systematic shifts in responses under role framing compared to the no-role baseline. The detective role—presented as the primary case in Section 5.1—showed the clearest lexical transitions, and the other four roles exhibited consistent domain-specific vocabulary increases under role framing (Section 5.1.4). Notably, user-side role assignment alone (Condition 2) altered response content even though the question sequence was identical, suggesting that models infer a different conversational contract when the user is framed as a professional actor. Shared user–AI role assignment (Condition 3) tended to amplify domain-consistent vocabulary and discourse patterns, and in representative cases (e.g., detective role), responses exhibited characteristics resembling peer-level professional exchanges (Sections 5.1–5.2). However, these findings should be interpreted as changes in generated discourse rather than validated changes in human learning or task outcomes.

6.2 Interpretation through Role Theory and Framing Theory

Role Theory Perspective. Role Theory conceptualizes roles as socially defined sets of expectations regarding goals, norms, and appropriate language use [8]. From this perspective, assigning a professional role via prompts can be understood as introducing an explicit set of role expectations into the interaction context. Our results are consistent with the idea that role expectations reorganize discourse toward role-consistent concerns and registers. For example, in the detective role, the same visual elements were increasingly interpreted in terms of investigation, risk, and prevention-related reasoning as role framing became stronger (Sections 5.1.2 and 5.2).

Importantly, the user-only condition (Condition 2) suggests a role-theoretic mechanism beyond “AI persona setting.” When the user is framed as a professional (e.g., “I am a detective”), the model may infer that the interlocutor possesses domain expectations and seeks expert-calibrated discourse, which can shift the style and depth of the response even without an explicit AI role instruction. The both-role condition (Condition 3) can be interpreted as establishing shared role expectations, enabling a more aligned exchange in which the model more readily deploys role-consistent vocabulary and professional concerns.

Framing Theory Perspective. Framing Theory explains how communication highlights particular aspects of a situation and thereby shapes interpretation and reasoning [9]. In our study, role instructions functioned as explicit frames that changed the interpretive axis applied to a constant stimulus. In the detective role, the scene was reframed from a “peaceful landscape” into a “potential incident context,” yielding hypotheses, inferred intentions, and risk-oriented considerations. This illustrates a framing effect: even with identical questions, role framing changes what is considered salient and what constitutes a plausible explanation (Sections 5.1–5.2).

The framing lens also clarifies why the both-role condition tended to produce richer role-consistent discourse: shared frames reduce ambiguity about the expected interpretation mode and encourage the model to elaborate within the frame (e.g., moving from description → hypothesis → procedural considerations), while the no-role condition tends to remain at the level of general description.

6.3 Two-Layer Interaction Model: External Frames × Model-Specific Tendencies

Beyond role framing itself, cross-model comparisons indicate that role effects interact with model-specific expression tendencies (Section 5.2). We conceptualize this as a two-layer structure: (1) an external interaction frame imposed by role assignment, which sets the direction of interpretation and the professional register; and (2) model-specific response tendencies, which shape how the frame is expressed (e.g., more procedural vs. more symbolic vs. more narrative styles).

For example, under the same detective frame, some models tended to emphasize practical investigation steps and verification points, while others emphasized symbolic codes or narrative dramatization. This suggests that prompt engineering outcomes are partly determined by an interaction between the designer’s external

framing choices and the model’s internal stylistic/interactional tendencies. Figure 2 summarizes this two-layer interaction model: Layer 1 (external role frame) determines the discourse direction—what interpretive axis, professional register, and social contract govern the exchange—while Layer 2 (model-specific tendencies) shapes how the frame is expressed through procedural, symbolic, structured, or narrative styles. The observed dialogue characteristics emerge from the interaction of these two layers.

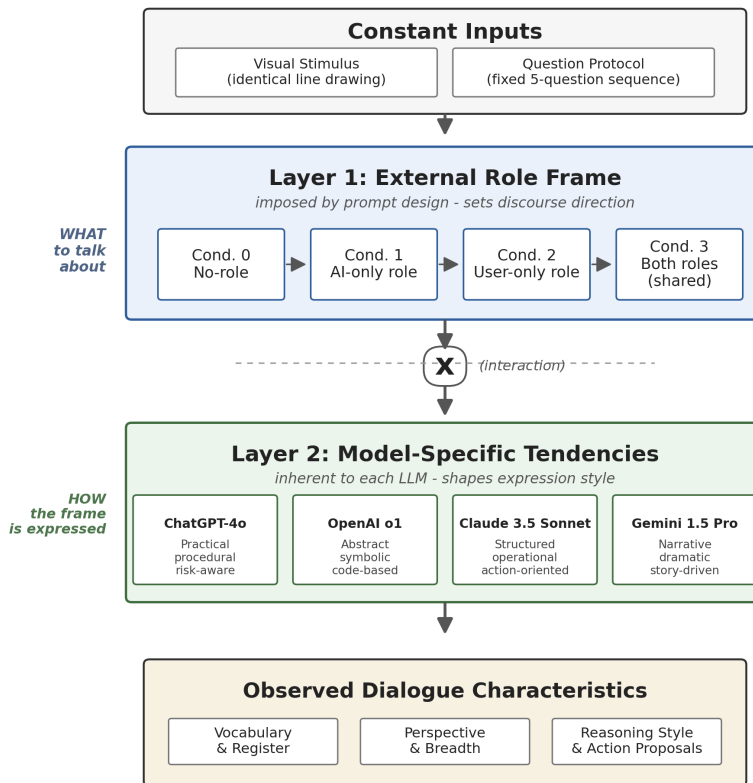


Figure 2. Two-layer interaction model of role-framed dialogue. Layer 1 (external role frame) determines the discourse direction—what interpretive axis, professional register, and social contract govern the exchange. Layer 2 (model-specific tendencies) shapes how the frame is expressed—through procedural, symbolic, structured, or narrative styles. The observed dialogue characteristics emerge from the interaction of these two layers. Representative examples are drawn from the detective role under the both-role condition (Section 5.2).

6.4 Implications for Prompt Engineering and Interaction Design

Recent surveys emphasize that prompt engineering comprises a family of techniques and that their effects are task- and context-dependent [2,3]. Our results suggest that

dual role assignment can be treated as a practical interaction-design lever, with at least three implications.

User-Side Role Assignment as a Dialogue-Depth Lever. In many real deployments, users already have professional identities (teacher, designer, engineer), but systems often treat them as neutral requesters. Our findings suggest that explicitly framing the user role (e.g., “As a teacher...,” “As a safety officer...”) can shift model responses toward expert-calibrated discourse even without changing the main question content. This may be particularly valuable in educational contexts where the goal is not a single “correct answer” but richer reasoning, perspective taking, or professional-style analysis [18].

Managing Role Persistence and Role Switching in Systems. In practical systems, users may change roles across tasks, and persistent memory may unintentionally carry role frames forward. To prevent unwanted role fixation while enabling role-based guidance, systems could benefit from explicit role-state management mechanisms, such as:

- a visible “current role state” indicator (user role / AI role) and a one-click role reset function,
- structured role templates (role name, goal, constraints, tone, evidence standard) to reduce ambiguity,
- explicit transitions when roles change (“Switch role: architect → safety inspector”), and
- safeguards for high-stakes contexts where role framing might encourage overconfident or speculative interpretations.

These design considerations are consistent with the broader prompt engineering view that interaction outcomes depend on how prompts are maintained across turns and tasks [2,3].

Beyond Visual Stimuli and Toward Multimodal Interaction. Although our study uses a simple line drawing as a controlled grounding artifact, the underlying mechanism—role frames shaping interpretation and discourse—should be applicable to other settings such as text-based problem solving, design reviews, and ethical reasoning dialogues. Moreover, as multimodal LLMs and image-grounded dialogue become increasingly relevant interaction paradigms [13], role framing may serve as a simple yet powerful means to steer multimodal conversations toward specific professional analyses. These extensions should be tested empirically, but the present study provides a controlled starting point.

6.5 Limitations

This study has several limitations that should be considered when interpreting the findings.

First, this is not a human-subject evaluation. The “user role” is implemented as prompt text executed by the operator, and we did not measure human perceptions,

learning outcomes, or task performance improvements. Therefore, claims about educational effectiveness or user benefits should be treated as implications to be validated in future user studies.

Second, the study is exploratory and descriptive in sampling. Each condition–role–model cell was sampled once within a limited time window; LLM outputs are known to be non-deterministic and can change with model updates. We mitigate this by releasing the full logs and final responses for audit and reanalysis (Section 4.7), but statistical generalization is not the aim of this paper.

Third, the quantitative analysis focuses on lexical frequency trends in Japanese text with English glosses for reporting (Section 4.5). Output length differences and translation choices may influence apparent lexical patterns. We mitigated the length confound by computing normalized domain-term densities per 1,000 characters (Section 5.1), which confirmed that both-role dominance in domain-specific vocabulary is not an artifact of response length. However, normalization was applied only to the detective role as a representative case; readers should consider length variation when interpreting other roles.

Finally, specialized vocabulary and “expert-like” framing do not guarantee correctness. Role framing may also increase the risk of plausible but ungrounded speculation in some contexts. Future work should incorporate correctness/safety evaluations alongside discourse-characteristic analysis.

6.6 Future Work

Future research should (i) conduct human-subject studies to evaluate perceived usefulness, learning outcomes, and collaboration quality; (ii) extend stimuli and tasks beyond a single image and include text-only and multimodal tasks; (iii) investigate role switching and persistence under memory-enabled systems; and (iv) develop and validate design patterns for role-state management in real applications. In addition, richer quantitative measures (e.g., length-normalized lexical metrics, discourse structure measures, and correctness/error analyses) could strengthen the evaluation of role-framing effects.

7 Conclusion

This paper examined how professional role assignment reshapes image-grounded dialogue with large language models (LLMs) when roles are assigned to the AI only, the user only, or both. Using an identical visual stimulus and a fixed five-question protocol, we compared four role-assignment conditions (no-role, AI-only role, user-only role, and both roles) across five professional roles and four contemporary LLMs (Section 4). We analyzed the resulting dialogue logs through quantitative term-frequency analysis and qualitative close readings (Section 5), and interpreted the observed patterns through Role Theory and Framing Theory (Section 6).

Three main findings emerged. First, role framing systematically altered lexical choices and interpretive focus, even under identical questions and a constant stimulus. In the representative detective case, responses shifted from neutral scene description

toward investigation- and risk-oriented discourse, including increased use of role-consistent terminology (Section 5.1). Second, user-side role assignment alone (Condition 2) changed response characteristics, suggesting that models adjust their discourse when the user is framed as a professional actor rather than a neutral requester (Sections 5.1–5.2). Third, shared role assignment between user and AI (Condition 3) often amplified domain-consistent vocabulary and discourse patterns, producing dialogues that exhibit characteristics resembling peer-level professional exchange, while still reflecting model-specific expression tendencies (Section 5.2).

These results should be interpreted within the scope of the study. The dialogues were generated under a controlled protocol executed by the author/operator, and the analysis focuses on changes in generated discourse rather than validated changes in human behavior, learning outcomes, or task performance. Moreover, the study is exploratory and descriptive in sampling (one run per condition–role–model cell within a limited time window), and LLM outputs may vary due to non-determinism and model updates (Section 4.7). Therefore, claims about educational effectiveness or user benefit should be treated as implications to be tested in future work rather than conclusions directly established by the present data.

From an interaction design perspective, the findings suggest that professional role assignment can be treated as an explicit interaction frame in prompt engineering, and that framing the user role is a practical lever for steering dialogue depth and perspective-taking. In addition, cross-model variation indicates that role framing interacts with model-specific response tendencies, supporting a two-layer view in which external frames shape “what to talk about,” while model traits shape “how the frame is expressed” (Section 6).

To support transparency and secondary analysis, all dialogue logs and extracted final responses (Japanese) are publicly available in the PROSPECT repository on Zenodo (DOI: 10.5281/zenodo.14567800) [14]. Future research should extend this work through (i) human-subject studies to evaluate perceived usefulness and learning outcomes, (ii) broader task and stimulus sets (including text-only and multimodal tasks), (iii) length-normalized and correctness-aware evaluations, and (iv) design investigations of role persistence, role switching, and role-state management in real-world systems.

Acknowledgments. The author acknowledges the use of generative AI tools (ChatGPT, Claude, Gemini) in this study, both as the subject of the research and as assistive tools for language refinement and structural organization during manuscript preparation. All content, analysis, interpretations, and final editorial decisions are the sole responsibility of the author.

Use of Generative AI Statement. In accordance with IxD&A policy on generative AI (<https://ixdea.org/ixda-policy-on-genai/>), the author declares the following uses of generative AI tools:

As research subjects. Four LLMs—ChatGPT-4o (OpenAI), OpenAI o1 (OpenAI), Claude 3.5 Sonnet (Anthropic), and Gemini 1.5 Pro (Google)—were used as the experimental subjects of this study for dialogue generation and analysis in Sections 4–5. All dialogue data were collected during December 5–10, 2024.

As manuscript preparation aids. During the preparation of this manuscript, the author used ChatGPT (GPT-4o, OpenAI) and Claude (Claude 3.5 Sonnet, Anthropic) for language refinement and structural organization in the initial draft of Sections 1–7; ChatGPT (GPT-5.2 Pro, OpenAI) and Claude (Claude Opus 4.6, Anthropic) for language refinement, structural reorganization, and revision support during the post-review revision of Sections 1–7; and Gemini (Google) for assistance in generating the conceptual diagram (Figure 2) in Section 6.3.

The author has reviewed and edited all AI-generated output and takes full responsibility for the accuracy and integrity of the content of this publication.

CRedit Author Statement: Keisuke Sato: Conceptualization, Methodology, Investigation, Formal Analysis, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project Administration.

Data Availability Statement. The complete dataset (dialogue logs and final responses) is publicly available in the PROSPECT repository on Zenodo: <https://doi.org/10.5281/zenodo.14567800>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. ASLERD and/or the member of the editorial board of IxD&A disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

References

1. Zhao, W.X., Zhou, K., Li, J., et al.: A Survey of Large Language Models. arXiv:2303.18223 (2023). doi:10.48550/arXiv.2303.18223.
2. Schulhoff, S., Ilie, M., Balepur, N., et al.: The Prompt Report: A Systematic Survey of Prompt Engineering Techniques. arXiv:2406.06608 (2024). doi:10.48550/arXiv.2406.06608.
3. Sahoo, P., Singh, A.K., Saha, S., et al.: A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications. arXiv:2402.07927 (2024). doi:10.48550/arXiv.2402.07927.
4. Bozkurt, A., Sharma, R.C.: Generative AI and Prompt Engineering: The Art of Whispering to Let the Genie Out of the Algorithmic World. Asian Journal of Distance Education, 18(2), i–vii (2023). doi:10.5281/zenodo.8174941.
5. Giray, L.: Prompt Engineering with ChatGPT: A Guide for Academic Writers. Annals of Biomedical Engineering, 51(12), 2629–2633 (2023). doi:10.1007/s10439-023-03272-4.
6. Kong, A., Zhao, S., Chen, H., et al.: Better Zero-Shot Reasoning with Role-Play Prompting. In: Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pp. 4099–4113 (2024). doi:10.18653/v1/2024.naacl-long.228.
7. Zheng, M., Pei, J., Logeswaran, L., Lee, M., Jurgens, D.: When “A Helpful Assistant” Is Not Really Helpful: Personas in System Prompts Do Not Improve Performances of Large

- Language Models. In: Findings of the Association for Computational Linguistics: EMNLP 2024, pp. 15126–15154 (2024). doi:10.18653/v1/2024.findings-emnlp.888.
8. Anglin, A.H., Kincaid, P.A., Short, J.C., Allen, D.G.: Role Theory Perspectives: Past, Present, and Future Applications of Role Theories in Management Research. *Journal of Management*, 48(6), 1469–1502 (2022). doi:10.1177/01492063221081442.
 9. Entman, R.M.: Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58 (1993). doi:10.1111/j.1460-2466.1993.tb01304.x.
 10. Nistor, N., Trăuşan-Matu, S., Dascălu, M., Duttweiler, H., Chiru, C., Baltes, B., Smeaton, G.: Finding student-centered open learning environments on the internet: Automated dialogue assessment in academic virtual communities of practice. *Computers in Human Behavior*, 47, 119–127 (2015). doi:10.1016/j.chb.2014.07.029.
 11. Fernandes, U. da S., Barbosa, G.A.R., Chagas, B.A., Barbosa, G.D.J., Barbosa, S.D.J., Prates, R.O.: Lessons Learned from Modeling the Interaction with Conversational Agents. *IxD&A*, No. 55, 139–173 (2022). doi:10.55612/s-5002-055-007.
 12. Scholten, M.R., Kelders, S.M., Van Gemert-Pijnen, J.E.W.C.: Similarity in Action with an Embodied Conversational Agent: Can Synchronous Speech Yield Higher Levels of Rapport? *IxD&A*, No. 56, 130–155 (2023). doi:10.55612/s-5002-056-007.
 13. Yin, S., Fu, C., Zhao, S., et al.: A survey on multimodal large language models. *National Science Review*, 11(12), nwae403 (2024). doi:10.1093/nsr/nwae403.
 14. Sato, K.: PROSPECT: Professional Role Effects on Specialized Perspective Enhancement in Conversational Task. *Zenodo* (2024). doi:10.5281/zenodo.14567800.
 15. Chen, B., Zhang, Z., Langrené, N., Zhu, S.: Unleashing the potential of prompt engineering for large language models. *Patterns*, 6(6), 101260 (2025). doi:10.1016/j.patter.2025.101260.
 16. Aslan, I., Neu, D., Neupert, D., Grafberger, S., Weise, N., Pfeil, P., Kuschewski, M.: How to Compliment a Human – Designing Affective and Well-being Promoting Conversational Things. *IxD&A*, No. 58, 157–184 (2023). doi:10.55612/s-5002-058-007.
 17. Lin, J., Tomlin, N., Andreas, J., Eisner, J.: Decision-Oriented Dialogue for Human-AI Collaboration. *Transactions of the Association for Computational Linguistics*, 12, 892–911 (2024). doi:10.1162/tacl_a_00679.
 18. Cain, W.: Prompting Change: Exploring Prompt Engineering in Large Language Model AI and Its Potential to Transform Education. *TechTrends*, 68, 47–57 (2024). doi:10.1007/s11528-023-00896-0.