2026 SIGdial Atlanta, GA, USA

Bridging Talk and Thought: Understanding Dialogue Dynamics Across Collaborative Problem-Solving Contexts

Zhengyuan Liu, Stella Xin Yin, Min-Yen Kan, Nancy F. Chen

Agency for Science, Technology and Research (A*STAR), Singapore
National University of Singapore
Nanyang Technological University, Singapore

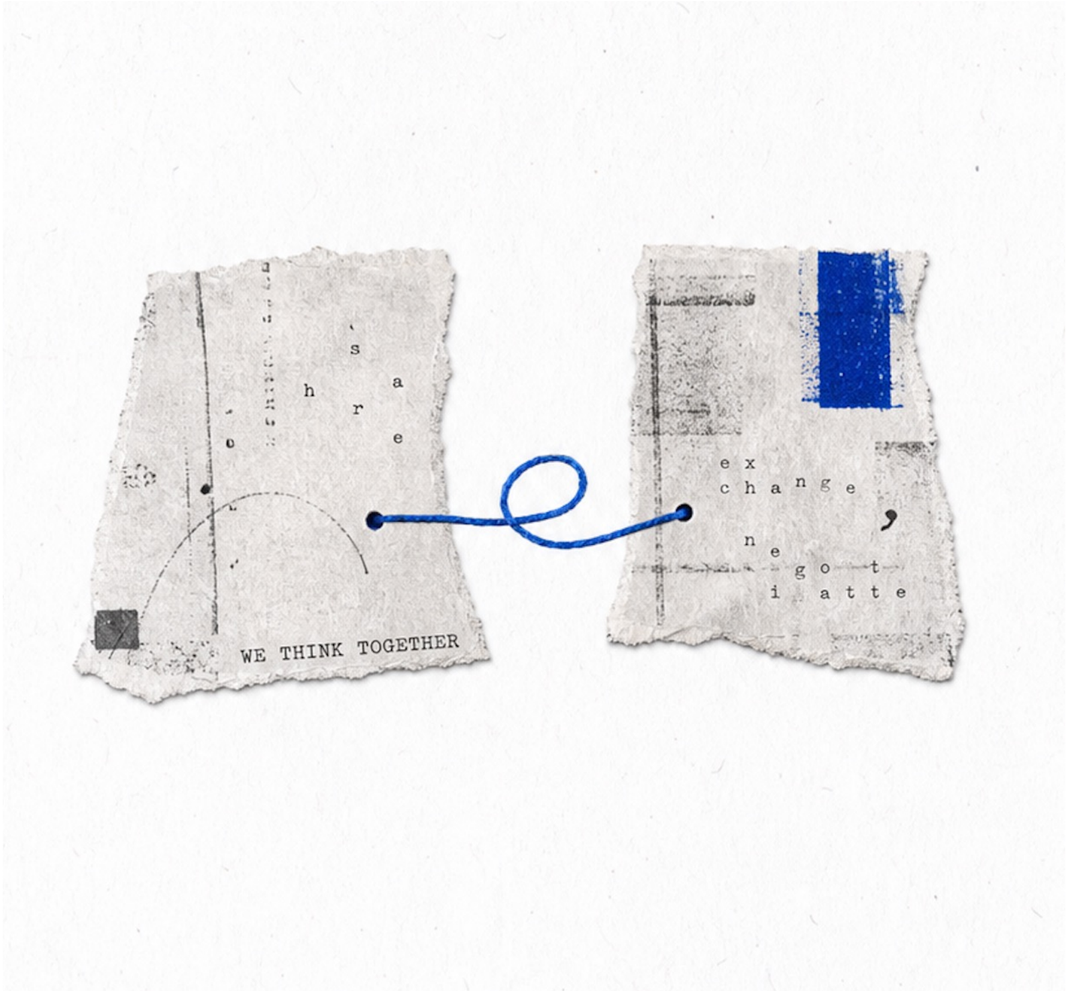
Agency for Science, Technology and Research SINGAPORE
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE
NUS National University of Singapore

BACKGROUND

Dialogue is where collaboration becomes observable

Collaborative problem solving is a collective process in which two or more agents share, exchange, and negotiate their **knowledge, skills, and efforts** toward a shared solution.

Through dialogue, collaborators externalize thinking, negotiate meaning, evaluate solutions, and regulate collective progress. Talk is both the vehicle and the catalyst of collaboration.



BACKGROUND

Human–AI work is moving to long horizons

CONVENTIONAL TOOL USE

Minutes, single session

One request, one response. Goals arrive fully specified; no progress to monitor and nothing to revise.

EMERGING COLLABORATION

Days or weeks, many sessions

Full-pipeline workflows: agents negotiate goals, track progress and revise strategy across sessions.

Shared understanding

Both parties must keep a common picture of the problem as it changes.

Sustained motivation

Socio-emotional work keeps engagement alive over long stretches.

Joint responsibility

Regulatory duties have to be carried by both sides, not one.

Human–AI pairs cooperate; and less collaborate

Across medical diagnosis and data analysis, interaction settles into **asymmetric tool use** : one party sets goals and monitors progress while the other responds reactively.

RELATED WORK

Three strands, still unconnected

STRAND 1

Dialogue modelling in NLP

Discourse acts, intent classification, state tracking and interaction-quality evaluation in task-oriented settings.

Asher et al. 2016 · Budzianowski et al. 2018 · Yin et al. 2025

STRAND 2

Learning sciences

Exploratory talk, socio-emotional regulation, and the self- / co- / shared-regulation triad; agents designed to prompt shared regulation.

Mercer & Wegerif 2002 · Hadwin et al. 2011, 2017 · Edwards et al. 2025 (BJET)

STRAND 3

Agent evaluation

Decision-oriented dialogue, multi-agent benchmarks with milestone KPIs, and taxonomies of coordination failure.

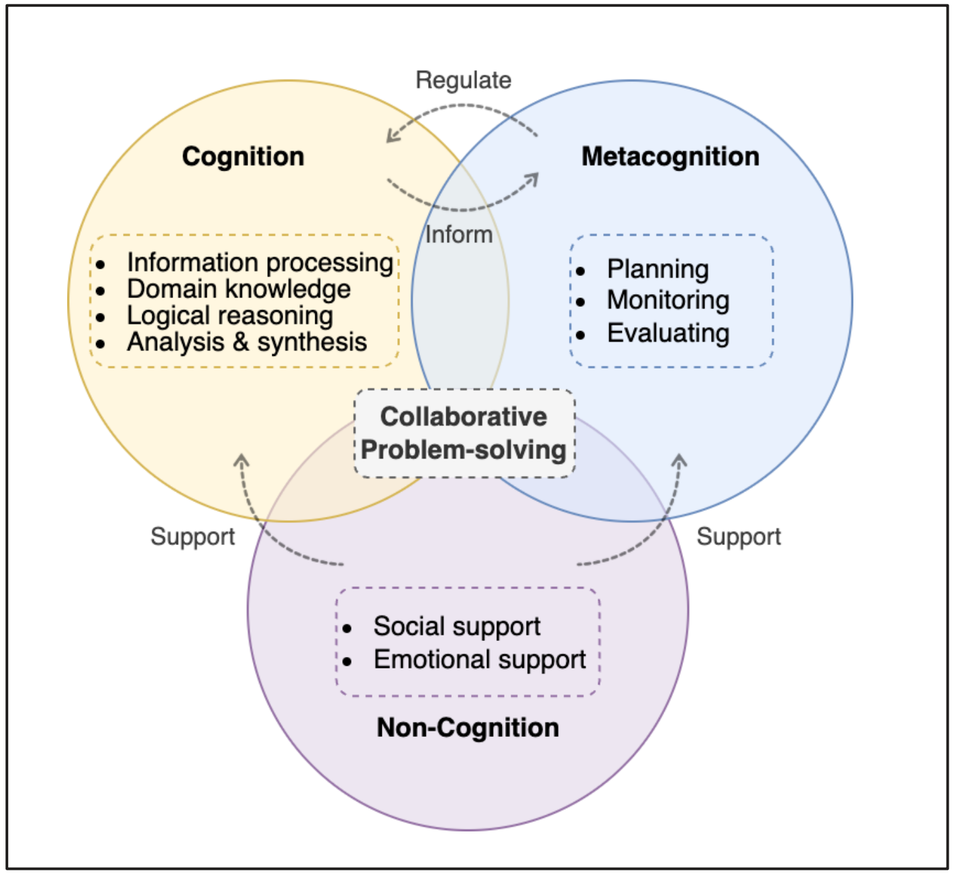
Lin et al. 2024 · Zhu et al. 2025 (MultiAgentBench) · Cemri et al. 2025 (MAST)

What existing analyses cannot see:

APPROACH	WHAT IT CAPTURES	BLIND SPOT
Dialogue acts & state tracking	Communicative intent and slot-level task state per turn	Who sets the goal, tracks progress and owns the plan
Outcome benchmarks	Task success and team-level KPIs under different topologies	The interaction process that produced the outcome
Failure taxonomies	Recurring coordination defects across execution traces	Early, continuous signals before a breakdown occurs
Regulation coding (SSRL)	Self-, co- and shared regulation in student groups	Comparability across domains, and human–AI dialogue

CONCEPTUAL MODEL

Three dimensions, one dynamic loop



Metacognition regulates

Planning, monitoring and evaluating — oversight of individual and collective thinking.

Cognition informs

Domain knowledge, reasoning, analysis and synthesis — the content regulation acts on.

Non-cognition supports

Social and emotional expression that sustains engagement and group cohesion.

Sustained human–AI work needs all three: shared understanding, motivation, and joint responsibility over time.

CODING SCHEME

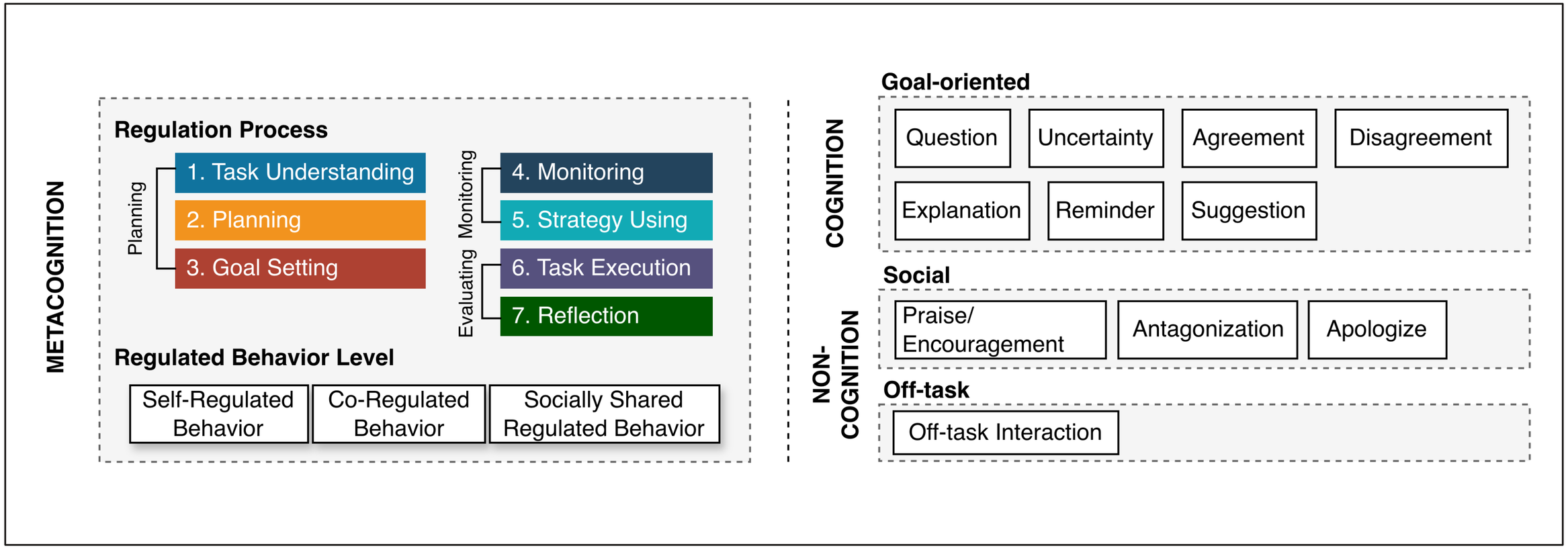
A hierarchical, two-layer annotation

Upper layer — dialogue & speaker level

Seven regulation processes; three regulated behavior levels (SR / CR / SSR).

Lower layer — utterance level

Seven goal-oriented, three social, and one off-task utterance type.



METACOGNITION · LAYER 1

Seven regulation processes

PHASE	PROCESS	GUIDING QUESTION IN THE CPS STAGE MODEL
Planning	① **Task Understanding**	What's going on? What do we know? What are the root causes?
	② **Planning**	What could we do? — team structure and task distribution
	③ **Goal Setting**	What's the best thing to do? — objectives and selection criteria
Monitoring	④ **Monitoring**	Tracking progress, summarising work done, naming what remains
	⑤ **Strategy Use**	How do we go about it? — proposing and evaluating solutions
Evaluating	⑥ **Task Execution**	Have we solved the problem? — implementing the chosen solution
	⑦ **Reflection**	Can we improve on what we have done? — adjustment and refinement

Groups revisit, skip and cycle through these processes; the sequence is **analytic, not linear**.

METACOGNITION · LAYER 1

Who drives the regulation?

Three behavior levels, distinguished by who initiates and maintains regulatory work.

INDIVIDUAL

Self-regulated (SR)

Regulating one's own cognition, motivation and behavior; owning a task component.

MARKER · “I”

> *“I'll track our progress on paper then.”*

INTERPERSONAL

Co-regulated (CR)

One member's regulation scaffolds or constrains another's; distributes cognitive load.

MARKER · “YOU”

> *“Who remembers every single turn?”*

COLLECTIVE

Socially shared (SSR)

Regulatory responsibility distributed across the group; jointly negotiated goals.

MARKER · “WE”

> *“What if we set a time limit for each section?”*

Over-reliance on SR causes coordination failure; excessive CR creates dependency. Effective CPS integrates all three.

METACOGNITION · LAYER 1

The two axes cross

Each regulation process can be enacted at any behavior level. Four of the seven, with annotated examples.

PROCESS	SELF-REGULATED (SR)	CO-REGULATED (CR)	SOCIALLY SHARED (SSR)
① **Task Understanding**	*“I need to figure out what this problem is asking for.”*	*“Do you agree that we need to understand what's being asked?”*	*“I think we should also check what resources are already available to us.”*
② **Planning**	*“I'm going to look at the data first.”*	*“Let's split up the work. Maria, could you handle the research part?”*	*“We could create a shared document to track progress. Does that work?”*
④ **Monitoring**	*“I've completed the first part of the analysis.”*	*“Have you finished the analysis section, Sam?”*	*“We still need the risk section. Should we reassign it or adjust the timeline?”*
⑦ **Reflection**	*“I think I can recreate the scatter plot to visualize the result.”*	*“Our solution didn't work. Should we try the alternative we discussed?”*	*“We overlooked the seasonal variations — Sarah mentioned this before, right?”*

Pronoun scope is the first cue, but coding is functional: the same words can be SR or SSR depending on what they regulate.

COGNITION & NON-COGNITION · LAYER 2

Eleven utterance types

Goal-oriented

Seven types of exploratory talk

- **QUE** Question
- **UNC** Uncertainty
- **AGR** Agreement
- **DIS** Disagreement
- **EXP** Explanation
- **REM** Reminder
- **SUG** Suggestion

Social

Socio-emotional climate

- **PE** Praise / Encourage
- **ANT** Antagonization
- **APO** Apologize

Off-task

Coded with context

- **OT** Off-task

Off-task talk can support the socio-affective side of collaboration, so we label each turn using the preceding and following utterances rather than the turn alone.

COGNITION & NON-COGNITION · LAYER 2

The definitions annotators work from

QUE Question

Inquiry seeking clarification, confirmation or more information about the task.

UNC Uncertainty

Doubt or low confidence about ideas or decisions, signalling a need for support.

AGR Agreement

Consensus with or acceptance of another's opinion, suggestion or idea.

DIS Disagreement

Dissent or misalignment, typically offering an alternative perspective.

EXP Explanation

Clarification, elaboration or reasoning that builds shared understanding.

REM Reminder

Redirecting attention to requirements, time limits or agreed goals.

SUG Suggestion

Proposing new ideas, approaches or solutions for the task.

PE Praise / Encourage

Positive feedback or validation that lifts morale and confidence.

ANT Antagonization

Harmful comments, conflict-seeking, criticism of personal attributes or irritation.

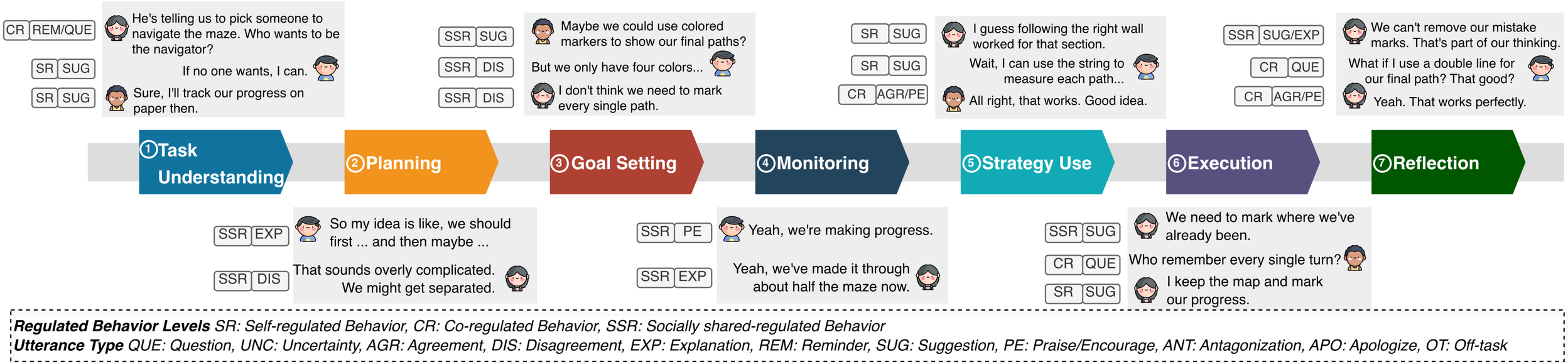
APO Apologize

Acknowledging a mistake or misunderstanding with an expression of regret.

WORKED EXAMPLE

The scheme applied to a maze-puzzle dialogue

Three students navigate a maze; the conversation moves from task understanding to reflection, with every utterance double-coded.



Regulated Behavior Levels *SR: Self-regulated Behavior, CR: Co-regulated Behavior, SSR: Socially shared-regulated Behavior*
Utterance Type *QUE: Question, UNC: Uncertainty, AGR: Agreement, DIS: Disagreement, EXP: Explanation, REM: Reminder, SUG: Suggestion, PE: Praise/Encourage, ANT: Antagonization, APO: Apologize, OT: Off-task*

CORPUS SELECTION

Nine CPS datasets across four contexts

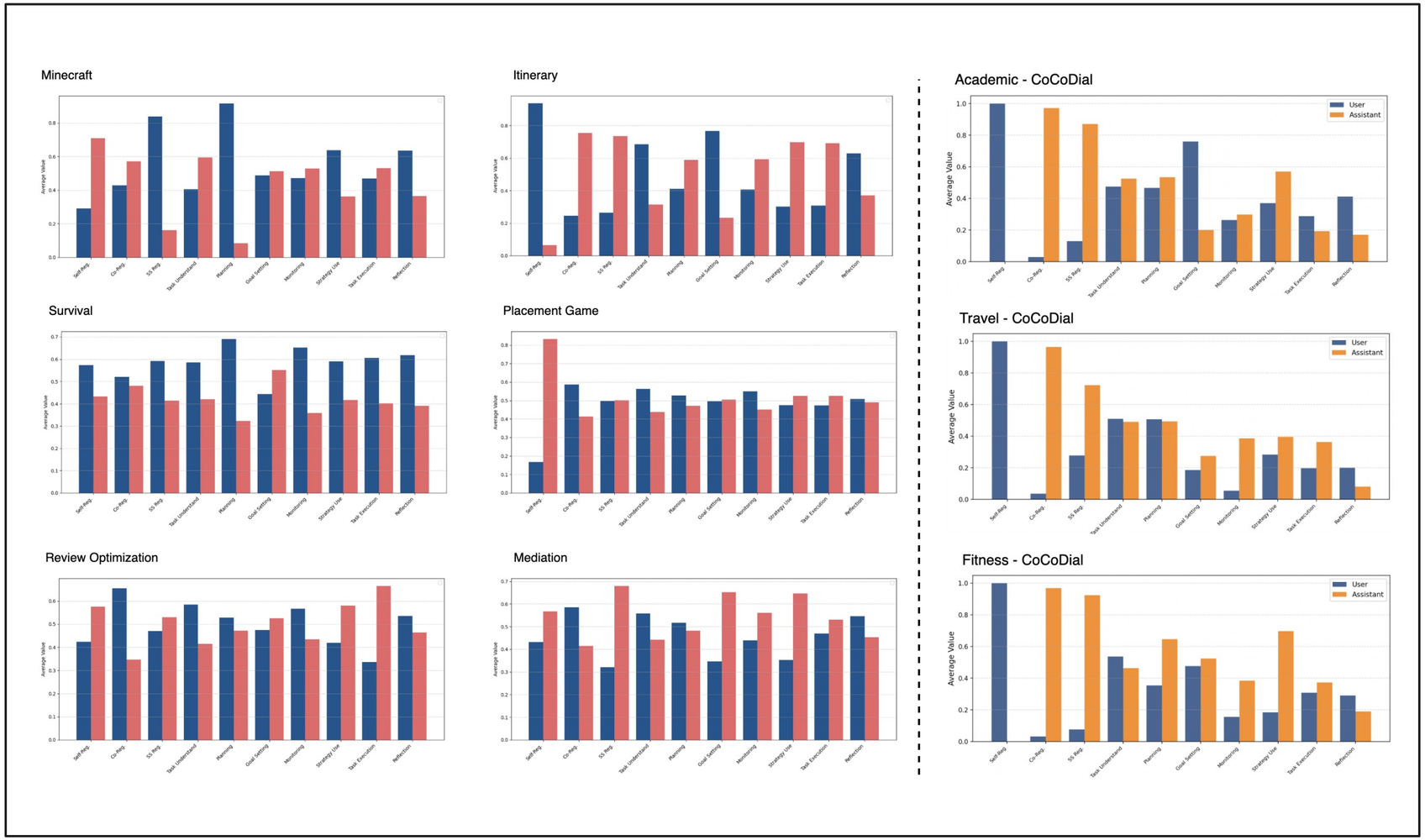
ID	DATASET	CONTEXT	TASK	PARTICIPANTS
D1	**GAP Corpus**	Survival scenario	Winter survival task — rank 15 items by importance	Human–human
D2	**Placement Game**	Game-based	2D object placement designed to elicit balanced collaboration	Human–human
D3–D5	**Decision-oriented dialogues**	Workplace	Review optimization, itinerary planning, travel mediation	Human–human
D6	**MSDC**	Game-based / build	Discourse-annotated collaborative building in Minecraft	Human–human
D7–D9	**CoCoDial**	Exploratory planning	Academic, travel and fitness planning with an assistant	**Human–AI**

Criteria: goal-directed problem solving, contextual diversity, and enough dialogue for meaningful analysis of dynamics.

FINDING 1

Regulatory balance depends on task structure

Structured environments induce functional asymmetry: collaboration patterns are context-dependent.



Balanced

Review Optimization and Mediation: even spread across all three dimensions.

Directive asymmetry

Minecraft and Itinerary: one partner leads Planning, Goal Setting and Execution.

Leadership without dominance

Placement Game: higher self-regulation for one speaker, equitable cognitive contributions.

FINDING 2

In human–AI dialogue, regulation splits almost cleanly

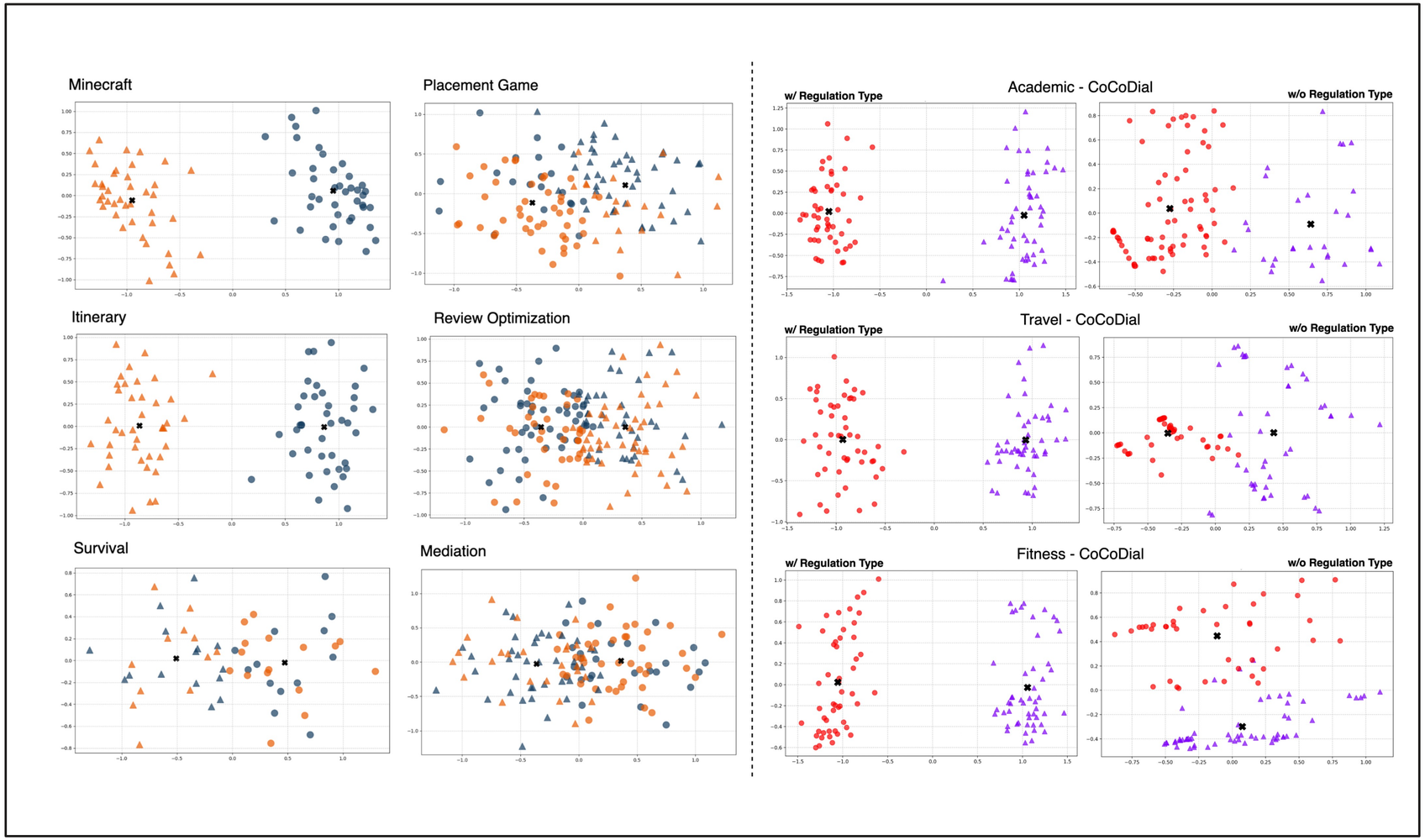
SELF-REGULATION	CO- & SHARED REGULATION	PROCESSES
Human ≈ 1.0 · Assistant ≈ 0.0	**Assistant dominates**	**Human leads the goal setting**
The user alone reports and manages their own state and intentions.	The assistant addresses and mirrors the user, but always after being prompted.	The assistant seldom initiates planning, goal setting or monitoring.

Because responses are always prompted rather than initiated, the human bears the full regulatory burden. **Detecting this asymmetry is the prerequisite** for systems that share regulation instead of awaiting instruction.

FINDING 3

Metacognitive features are what separate roles

K-means clustering of speaker vectors. Colors mark explicit roles; shapes mark roles that emerge in the dialogue.



Structured tasks split

Minecraft and Itinerary: two clearly separated speaker clusters.

Negotiation overlaps

Mediation and Placement Game: overlapping clusters, symmetrical roles.

Ablation on CoCoDial

With regulation types: tight, distinct clusters. Without them: mixed, diffuse distributions.

Regulation behaviors are essential discriminators for modeling collaboration.

CONCLUSION

Three takeaways

01 Dialogue makes collaboration measurable

Metacognition, cognition and non-cognition are readable turn by turn from the conversations

02 Regulation can be a discriminating signal for human-AI collaboration

Remove metacognitive features and speaker roles might collapse into indistinct clusters.

03 Balance depends on task structure

Structured tasks induce functional asymmetry; negotiation-heavy tasks stay more symmetrical.

Thank you for your attention! Questions are welcome!