

Planning Ahead with RSA: Efficient Signalling in Dynamic Environments by Projecting User Awareness across Future Timesteps

Anwasha Das
Saarland University
Saarbrücken, Germany
adas@lst.uni-saarland.de

Jörg Hoffmann
Saarland University
Saarbrücken, Germany
hoffmann@cs.uni-saarland.de

John Duff
UCLA
Los Angeles, USA
duff@ucla.edu

Vera Demberg
Saarland University
Saarbrücken, Germany
vera@lst.uni-saarland.de

ABSTRACT

Adaptive agent design offers a way to improve human-AI collaboration on time-sensitive tasks in rapidly changing environments. In such cases, to ensure the human maintains an accurate understanding of critical task elements, an assistive agent must not only identify the highest priority information but also estimate how and when this information can be communicated most effectively, given that human attention represents a zero-sum cognitive resource where focus on one message diminishes awareness of other or upcoming information. We introduce a theoretical framework for adaptive signalling which meets these challenges by using principles of rational communication, formalised as Bayesian reference resolution using the Rational Speech Act (RSA) modelling framework, to plan a sequence of messages which optimise timely alignment between user belief and a dynamic environment. The agent adapts message specificity and timing to the particulars of a user and scenario based on projections of how prior-guided interpretation of messages will influence attention to the interface and subsequent belief update, across several timesteps out to a fixed horizon. In a comparison to baseline methods, we show that this effectiveness depends crucially on combining multi-step planning with a realistic model of user awareness. As the first application of RSA for communication in a dynamic environment, and for human-AI interaction in general, we establish theoretical foundations for pragmatic communication in human-agent teams, highlighting how insights from cognitive science can be capitalised to inform the design of assistive agents.

KEYWORDS

Pragmatic Reasoning; Human-Aware Planning; Rational Speech Acts Framework; AI Assistance; Adaptive Agents

ACM Reference Format:

Anwasha Das, John Duff, Jörg Hoffmann, and Vera Demberg. 2026. Planning Ahead with RSA: Efficient Signalling in Dynamic Environments by

Projecting User Awareness across Future Timesteps. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 9 pages. <https://doi.org/10.65109/BRGK2754>

1 INTRODUCTION

In high-stakes environments, multiple critical events often unfold simultaneously, making it difficult for human overseers to maintain situational awareness of rapidly evolving scenes. Consider an air traffic controller managing several aircraft during a storm, or a power grid control room, as illustrations. In such settings where human perceptual lapses (due to overload and eroding situational awareness) carry high risk, assistive AI agents can support operators who must multitask under increased cognitive demands to sustain effective oversight [4, 40].

The efficacy of AI assistance, however, is limited by the fact that human attentional resources are finite. Once a message captures and engages attention, it temporarily constrains an individual’s ability to perceive and process new information [43]. In dynamic settings with multiple concurrent events, the timing and content of communications influence whether human partners maintain situational awareness or experience overload [12]. Poorly-calibrated communication has severe consequences in such safety-critical situations (see 1, 19 and 33 for documented accidents). A pattern follows: although today’s intelligent systems are adept at identifying all important (or critical) information quickly amidst vast data streams, they do not yet (sufficiently) take into account human cognition when communicating this information [36].

The true value of AI assistance, hence, comes about not from indiscriminately alerting human partners of all critical information but depends instead on the assistive agent’s capacity to deliver alerts that are timed and framed to align with individual operators’ evolving mental model (beliefs and attentional resources) of the world [8, 16].

Suppose you are an operator who must oversee a swarm of fire-fighting drones – a dynamic, high-stakes environment where time is constrained. When concurrent risks emerge: (a) drone 2 facing high wind-speed conditions, shortly followed by (b) drone 3’s imminent battery failure, each requires urgent but distinct alert messages from the AI system. It would be prudent for a good assistive system



This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). <https://doi.org/10.65109/BRGK2754>

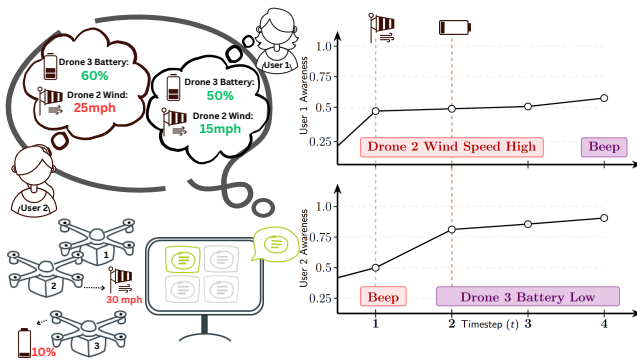


Figure 1: Impact of individualised alert sequences (x-axis) for distinct users in the same environment on user awareness (y-axis). Timesteps (x-axis) are annotated with the onset of critical events. Our model uses recursive Bayesian reasoning to construct user-specific models of how each user updates their world beliefs in response to alert sequences, optimising for global situational awareness over time.

to understand you (your prior beliefs, attention, and environment) to be able to choose a sequence of alerts which would most quickly lead you to a correct awareness of these risks [8]. For example, in this scenario (shown in Fig. 1), one solution would be to deliver a precise alert about the critically high wind-speed before moving to the battery issue, although this may delay the timing of the latter alert. For some users (e.g., User 1 in Fig. 1), this would be a good choice, as both issues must be dealt with, and attempting a shorter, less specific message about the wind could leave a user confused. However, if you (User 2 in Fig. 1) were already anticipating the high wind-speed, but are unaware of the battery issue, an AI assistant should take advantage of this existing knowledge: it could choose the shorter wind-speed message, knowing it would be correctly interpreted, and thus allow for a more timely notification about the unexpected battery failure.

Our work develops a theoretical framework for optimising the delivery of alerts (or messages) that takes into account the human partner’s attention and awareness when choosing how and when to communicate critical information. By reasoning about both the world and the human’s evolving mental model of it, we aim to enable assistive AI agents that can generate messages that are truly context and human-aware. As Miehling et al. [32] attest, AI agents built atop cognitive models of rational human communication—capable of reasoning about human partners’ beliefs, goals, and attentional limits—can enable uniquely sophisticated partnerships. This capability grows increasingly necessary as the AI we interface with expands out from constrained, screen-based interactions to more complex, high-stakes, and dynamic environments: (semi-)autonomous vehicles navigating complicated urban traffic [27], robotic surgeons adapting to complications [41], or drones coordinating disaster responses [25].

First, to model rational communication, we draw from research in pragmatics, in particular the Rational Speech Acts Framework (RSA) [17], a computational framework which uses Bayesian reasoning to model interactions between speakers and listeners. It captures how

speakers tailor messages by recursively simulating the listener’s beliefs about messages they hear, allowing RSA-based approaches to enrich data-driven models on tasks like scene description, referential expression generation and instruction following [3, 15, 35]. However, vanilla RSA lacks multi-step reasoning and only optimises for single-turn utterances (or alerts), making it ill-suited for dynamic monitoring tasks that require incremental, adaptive communication. For instance, vanilla RSA myopically optimises current alerts (like User 1’s wind-speed alert) without considering their time impact on future communications. Even though neural models offer alternatives, they produce overly verbose outputs violating human pragmatic preferences – problematic when communicative efficiency is key [31, 42].

To address this gap, we propose a pragmatic signalling framework that extends RSA with finite-horizon lookahead planning. Our approach preserves RSA’s strengths of recursive communication modelling while extending it to real-world settings through reasoning about sequences of messages over multiple timesteps. Thus, it models (a) the human’s evolving estimated beliefs, (b) the dynamics of the environment, and (c) the projected communicative utility of different alert sequences. By estimating dynamic, user-specific belief updates up to a planning horizon, our system strategically decides *what*, *how*, and *when* to communicate to the human partner, ensuring that messages are conducive and contributive to the operator’s pre-existing as well as future knowledge and attention. This adaptive, context-sensitive approach, grounded in conversational efficiency, establishes theoretical foundations for assistive agents that can reason about the temporal dynamics of human attention and belief evolution in changing environments.

2 PROBABILISTIC PRAGMATIC MODELS

The Rational Speech Act (RSA) framework [11, 13, 17] formalises effective communication as Bayesian optimization, where conversational partners reason recursively about each other’s mental states. Pragmatic behaviour, moving beyond literal communication alone, arises naturally from this recursive social inference process. Literal meanings are enriched through strategic inferences from prior knowledge and balance between informativity and cost. Classic RSA models can explain phenomena like scalar implicature (‘some’ implying ‘not all’) [14, 18] and ambiguous reference resolution [13, 20, 34] by simulating how listeners infer speaker intentions, and vice versa. This has allowed RSA to be successfully applied in diverse settings from reference games [13] to grounded visual scene descriptions [6]. However, vanilla RSA’s successes are fundamentally limited: the framework is optimised for single-turn messages, useful only in static environments. So, while RSA determines *how* to convey important information in isolation, it cannot plan *what* to prioritize or *when* to intervene as environments evolve—a consequential limitation in dynamic, multi-step settings, like ours.

To address these limitations, previous work has pursued partial extensions toward temporal reasoning. Cohn-Gordon et al. [7] extend RSA to score partial messages based on a projected listener interpretations of probable continuations. Their *incremental RSA* model prioritizes packing information as early in the message as possible, but applies this only to the next timestep utility of

these partial prefixes—appropriate perhaps for modelling human behaviour, but insufficient for global sequence efficiency.

To achieve our goals, we develop an alternative RSA extension, which also prioritizes temporal efficiency within a multi-timestep sequence, but does so by considering finite-horizon projections of future message sequences. To estimate projected sequence efficacy, it is necessary to project how incremental interpretation affects user priors at future timesteps. This dynamic prior accumulation has been previously considered within RSA for multi-turn conversations [2]. We extend it here for a practical setting, and show how dynamic representation of user belief can be deployed within RSA as part of multi-step planning in an environment which is itself changing over time.

3 PROBLEM FORMULATION

We formalise the setting described above as a pragmatic assistive signalling framework, where an AI system (Speaker, S_2) communicates using utterances (messages) with the human operator (Listener, L_1) to help them track the state of the world over multiple timesteps (finite horizon $1 \leq t \leq H$). This interaction is described by the tuple $(S, \mathcal{P}, U, H, B, R)$, where S is the set of true world states, \mathcal{P} the set of properties constituting S , U a space of utterances the system can use to signal at time t , H the planning horizon, i.e., the maximum number of timesteps considered when selecting utterance sequences, B the distribution of the human operator’s beliefs, and R the reward function.

The world state S evolves dynamically, with each timestep t yielding $s_t : \mathcal{P} \mapsto V$ that assigns values to all properties $p \in \mathcal{P}$ (e.g., $s_t(\text{drone 4 battery level}) = 80\%$). While the system S_2 has full access to the true world state, the human L_1 , however, cannot directly observe the complete s_t . Instead, their beliefs about the world b_t are a noisy and lagging estimate of likely states, following changes in S only through observations of individual p . Thus, due to distractions or perceptual overload, L_1 often has outdated, false beliefs. This results in low situational awareness, which can be risky, especially if there are critical properties $Cr_t(p) = 1$. Certain properties become critical when $s_t(p)$ takes a value in a predefined critical (risk) region $StatCr(p) \subseteq V$, triggering $Cr_t(p) = 1$ (e.g., $StatCr(\text{drone 4 battery level}) = [0, 15\%]$ or $StatCr(\text{drone 4 rotor condition}) = [off]$). Formally, $Cr_t(p) = 1[s_t(p) \in StatCr(p)]$.

To help L_1 sustain high situational awareness and to communicate risks, S_2 selects utterances $u \in U$ from a predefined, finite lexicon \mathcal{M} where each u maps to a subset of \mathcal{P} via $\mathcal{M} : U \mapsto 2^{\mathcal{P}}$, $\mathcal{M} \in \{0, 1\}^{|U| \times |\mathcal{P}|}$. To illustrate, consider a simple lexicon \mathcal{M} consisting 3 utterances of various specificities: “Drone 4 battery level is low at 10%”, “Battery Low”, and “Beep”, all of which might direct attention to $Cr_t(p = \text{drone 4 battery level}) = 1$. Ergo, matrix \mathcal{M} tracks the (u, p) mapping permitting variations in both informativeness and length, when referring to a property p .

The assistive system S_2 ’s objective is to generate a sequence of utterances (u_1, \dots, u_H) that maximises reward r_t . The reward function R incorporates the L_1 ’s estimated belief distributions B_t to calculate the effectiveness of each utterance sequence in maximising belief alignment ($b_t \approx s_t$). The R can be formulated based on task or domain-specific goals—for our purposes, the goal is sustaining situational awareness about critical properties or risks.

4 MODELS

Let us now demonstrate how we can approach this problem using and extending RSA. We start with the Vanilla RSA agent and incrementally build in (i) belief updates, (ii) user-specific priors, and (iii) finite-horizon planning to help accomplish the task of pragmatic assistive signalling in dynamic environments.

4.1 Vanilla RSA: Modelling Pragmatic Reasoning

A **Literal Speaker** S_0 is one which stochastically selects utterances based solely on whether u literally means (or maps to) property p , without modelling the listener’s beliefs or informational needs. This is a naive interface. Using lexicon \mathcal{M} , the conditional probability $S_0(u | p)$ is split equally between every u whose meaning contains the target p :

$$S_0(u | p) \propto \frac{\exp(\log P(p | u))}{\sum_{u'} \exp(\log P(p | u'))} \quad (1)$$

Here, $P(p | u)$ is derived from the lexicon semantics: $P(p | u) \propto 1[\mathcal{M}(u, p) = 1]$, yielding a uniform distribution over utterances that refer to p .

At a level higher, the **Pragmatic Listener** L_1 , reasons about the speaker’s communicative intent. Given utterance u , L_1 infers which property p a S_0 speaker would most likely have intended to refer to, based on the speaker’s distribution $S_0(u | p)$ and their own prior belief $P(p) = \frac{1}{|\mathcal{P}|}$, $\forall p \in \mathcal{P}$:

$$L_1(p | u) = \frac{S_0(u | p) \cdot P(p)}{\sum_{p'} S_0(u | p') \cdot P(p')} \quad (2)$$

The top layer, **Pragmatic Speaker** S_2 , extends this nested reasoning to select utterances that most effectively convey the intended property p to L_1 . S_2 aims to minimise the listener’s uncertainty (or surprisal) about the intended meaning: $-\log L_1(p | u)$, adjusted by utterance costs (e.g., length, or complexity). Even though RSA allows an unbounded hierarchy of nested agents ($S_n(L_{n-1})$), studies suggest that a depth of two sufficiently captures the benefits of pragmatic reasoning in most communicative settings [44].

Nonetheless, vanilla RSA presents two fundamental limitations that make it insufficient for efficient signalling in dynamic environments: (1) it operates on a single utterance in isolation, lacking explicit models of how listener beliefs evolve and propagate across multiple steps; and (2) it is limited to the selection of an optimal message given a target property p as the goal. That is, while vanilla RSA can reason about a listener’s interpretation of utterances $u \in U$ at current time t , it cannot track how each u revises their beliefs and priors for subsequent timesteps. And because it provides no principled mechanism for determining which $p \in \mathcal{P}$ merits the listener’s attention, especially when multiple p may be simultaneously critical, the decision of *what* to communicate remains unresolved. One could repetitively invoke RSA at each timestep, or one could apply heuristic property selection, but such approaches fundamentally sidestep the core challenge: in dynamic environments where context and relevance evolve rapidly, intelligent assistive systems must not only utter a maximally pragmatic signal for current t but also deliberately determine which property of the evolving system deserves attention, and when.

4.2 Baseline Model: Tracking Beliefs over Time

We address vanilla RSA’s limitations by introducing temporal belief updates and reward-based property selection to modify the pragmatic speaker S_2 . The S_0 and L_1 agents remain as defined in Equations 1 and 2, but now operate with the assumption that the system has access to previous beliefs b_{t-1} and world context S (see Problem Formulation).

4.2.1 Belief Tracking over Time. The motivation for extending RSA lies in its Bayesian structure, which naturally supports adaptive approaches. In Bayesian models, the posterior from one timestep becomes the prior for the next – ‘today’s posterior is tomorrow’s prior’ [28] – lending itself well to tracking how user beliefs evolve in response to sequences of utterances, requiring temporal indexing across timesteps.

We operationalise how a user’s attention evolves after an alert at current time t given the updated L_1 distribution. Function AT maps the L_1 to an updated attention distribution as $AT : (L_1, at) \mapsto at$, where, for each property $p \in \mathcal{P}$, the updated attention $at_t(p)$ reflects the likelihood of the user attending to or understanding property p after observing utterance u_t . Formally, $\forall p \in \mathcal{P}, \forall t > 0$:

$$at_t(p_t | u_t) := AT[L_1(p_t | u_t; b_{t-1})] \quad (3)$$

Next, function B estimates how a user’s belief state b_t evolves given the prior belief b_{t-1} , current system state s_t , and attention at_t .¹ For each property $p \in \mathcal{P}$ and value $v \in \mathcal{V}$, the updated belief ($\forall t > 0$) is:

$$b_t(u_t | p_t) = B(b_{t-1}, s_t, at_t)(p)(v) := at_t(p) \cdot P(v = s_t(p)) + (1 - at_t(p)) \cdot b_{t-1}(p)(v) \quad (4)$$

where $P(v = s_t(p)) = 1$ if v is the true value of p at t , and 0 otherwise. High attention ($at_t(p) \rightarrow 1$) leads to stronger alignment with the true system value $s_t(p)$, while low attention ($at_t(p) \rightarrow 0$) preserves the prior belief. This formulation allows the model to track how each potential utterance $u \in U$ would influence user beliefs.

4.2.2 Rewards. To determine which property deserves attention based on overall situational awareness and task goals – we define a reward function that quantifies the communicative value of each utterance u_t in terms of its effect on belief alignment b_t with the true system state s_t , weighted by the task-relevance of each p – here, whether p_t is critical:

$$R_t(u_t | s_t, Cr_t) = \sum_{p \in \mathcal{P}} b_t^{u_t}(p)(s_t(p)) \cdot Cr_t(p), \quad \forall t > 0 \quad (5)$$

Intuitively, the reward is a scalar value that measures how well a u_t guides the listener’s mental model toward an accurate understanding of currently critical properties ($Cr_t(p) = 1$).

4.2.3 Pragmatic Speaker. The baseline S_2 now can make more informed u selections which account for both belief evolution and property prioritisation. The optimal utterance at each time $t > 0$ is then chosen as:

¹Since real user beliefs are not directly observable, B serves as a computational approximation. More details about alternate variants of functions (AT, B, R) are in the Supplementary Materials.

$$S_2^{\text{baseline}}(u_t | s_t, Cr_t) = \arg \max_{u \in U} R_t(u | s_t, Cr_t) \quad (6)$$

The baseline model represents our closest analogue to prior work, now extending vanilla RSA to dynamic environments by building in temporal belief tracking. While it can reason about evolving beliefs, it makes simplifying assumptions: uniform user priors and greedy (single-timestep) temporal reasoning. As such, it serves as the foundation for evaluating more sophisticated variants that we introduce next. We call it **Dynamic RSA (d-RSA)** for brevity.

4.3 Pragmatic Human-Aware Planning

Having established the baseline model, we introduce two further modifications that improve its adaptive pragmatic signalling ability: user priors and temporal planning. We establish these reasoning capabilities independently before combining them together to develop the full model.

4.3.1 Reasoning about Users: The baseline model’s uniform priors $P(p) = \frac{1}{|\mathcal{P}|}$ reflect a generic user assumption, treating all listeners as identical entities. Yet real users are diverse – bringing varied beliefs, biases and contextual awareness that influence how they interpret communicative signals significantly. We capture this heterogeneity by substituting these uniform priors with user-specific priors from the beginning: conditioning the L_1 listener on their previous belief state b_{t-1} and task-relevant contextual knowledge: $P(p | b_{t-1}, StatCr)$.

$$L_1(p | u, b_{t-1}, StatCr) = \frac{S_0(u | p) \cdot P(p | b_{t-1}, StatCr)}{\sum_{p'} S_0(u | p') \cdot P(p' | b_{t-1}, StatCr)} \quad (7)$$

These user-specific priors empower the model to reason about distinct user profiles (noisy users), who may, for e.g., maintain high prior expectations for familiar system properties (e.g., battery), or exhibit attention biases toward aspects they perceive as safety-critical based on prior knowledge (e.g., windspeed).

4.3.2 Reasoning about Time: The myopic approaches we’ve presented above—whether baseline or user-specific—optimise each utterance in isolation, sacrificing long-term goals for immediate rewards. To surmount this, we can extend the Pragmatic Speaker S_2 with finite-horizon planning. The planner generates all possible utterance sequences $\mathbf{u}_{1:H} = (u_1, u_2, \dots, u_H)$ up to planning horizon H , subject to duration constraints. For each complete sequence, it computes the total cumulative reward:

$$\mathcal{R}(\mathbf{u}_{1:H}) = \sum_{t=1}^H R_t(u_t | s_t, Cr_t) \quad (8)$$

subject to:

$$dur(u_t) = k_{u_t} > 1 \implies u_{t+1}, \dots, u_{t+k-1} \text{ are } \langle X \rangle$$

This constraint encodes temporal opportunity costs – longer utterances lock the system into a multi-step delivery, during which only $\langle X \rangle$ messages are legal. These $\langle X \rangle$ utterances assign zero attention to all p , representing cases where a listener’s attentional resources are already committed to process an ongoing multi-step utterance ($duration(u_t) = k(u_t) > 1$ timestep). They yield no belief updates and thus low or zero rewards, while also delaying other

reward-accruing utterances. This double penalty allows the planner to simulate trade-offs between informativeness and timeliness.

Applied with only uniform priors, the result of this user-agnostic planning would likely result in a less sensitive model. We use this deliberate ablation to isolate the contribution of planning (reasoning about time), terming it the **d-RSA + Planning** model. We term the model with informed priors (Eq. 7) but without planning the **d-RSA + Priors** model, to assess the impact of user-modelling (reasoning about users) individually.

4.3.3 Reasoning about Users and Time: Finally, combining user-specific priors (Eq. 7) with finite-horizon planning lets the agent track evolving user beliefs and evaluate complete utterance sequences. This yields our full model, **d-RSA + Priors + Planning**. The optimal utterance sequence U^* is the one that maximises cumulative reward over all candidate sequences $\mathbf{u}_{1:H}$:

$$U^* = \arg \max_{\mathbf{u}_{1:H}} \sum_{t=1}^H R_t(u_t \mid s_t, Cr_t) \quad (9)$$

The pragmatic speaker then executes the t^{th} element of U^* at each timestep $\forall t \in \{1, \dots, H\}$ as $S_2^{\text{full}}(u_t) = (U^*)_t$.

5 EXPERIMENTS

5.1 Setup

5.1.1 Environment. To evaluate our model’s adaptability across diverse challenging scenarios, we created 800 trials in a simulated Drone World environment. Each trial involves 4 drones with 6 attributes each (24 properties) with their values evolving over 7 timesteps; properties are either continuous (e.g., altitude) or binary (e.g., rotor condition). The environment simulates time-critical monitoring in safety-critical operations (e.g., emergency response), where a human overseer must maintain situational awareness of all 4 agents’ states to prevent failures (e.g., crashes, resource depletion).

Similar to the toy example in Problem Formulation, our lexicon \mathcal{M} contains utterances of various specificities to refer to each property. We define a duration vector $Dur \in \mathbb{N}^{|U|}$, where $Dur_u = k(u)$ denotes the timesteps required to utter u . For example, a specific message like ‘Drone 4 battery level is low’ consumes 3 timesteps, while a brief ‘Beep’ takes only 1. Together, $(\mathcal{M} + k(u))$ defines our lexicon, encoding both (u, p) mapping and temporal cost. We also introduce a special *informationless* utterance [39]: *Silence ‘...’* which distributes minimal attention uniformly across all p and not only helps capture naturalistic fluctuations of attention (due to scanning, mind-wandering, distractions, etc.) but more importantly provides an alternative to constant alerting when it is not optimal.

We systematically vary trial complexity as follows.

Critical Properties (hereafter, $Cr(p)$): Each scenario has 2 – 4 $Cr(p)$, sampled randomly from the full set of 24. Some trials have overlapping $Cr(p)$ features—multiple risks associated with the same drone or attribute, which we quantify with *overlap metric*: the proportion of shared features.

Temporal Dispersion: Each p becomes critical at specific timestep t . We vary the spacing between them between 0 – 3, with the first onset occurring $1 \leq t \leq 3$.

User Belief Modelling: The L_1 begins with an initial belief $b_{t=0}(p, v)$,

a probability distribution over each property p and its possible values v . We simulate diverse (misspecified or noisy) users by varying (i) the accuracy of $b_{t=0}(p, v)$ about critical properties (10%-80% correct) and (ii) general situational awareness (low vs. high accuracy and precision for non-critical properties).²

5.1.2 Implementation. Our approach employs finite-horizon planning with the objective to maximise (non-discounted) rewards, through a reward designed to integrate RSA with lookahead. For our experiments to fairly evaluate the model design variants, we need to know the optimal plan for each trial and its maximal attainable reward. To this end, we use breadth-first search to enumerate utterance sequences up to horizon $H = 7$, enforcing multi-timestep constraints to construct a pruned search tree.

5.1.3 Analysis. We evaluate the overall performance of the full d-RSA + Priors + Planning model and its two key capabilities:

- (A) **Multi-step Planning:** Does planning utterance sequences based on their project rewards over a finite horizon, yield better communicative outcomes than selecting the optimal option myopically at every free timestep?
- (B) **Belief Tracking:** Does selecting utterance sequences by reasoning about the users’ estimated beliefs improve signalling compared to assuming generic, uniform priors?

To recall, our analysis compares four models: the baseline Dynamic RSA (d-RSA), which extends vanilla-RSA to a dynamic environment but is greedy and user-agnostic; d-RSA + Priors, which adds user-modelling but remains greedy; d-RSA + Planning, which adds planning but uses uniform priors; and the full d-RSA + Priors + Planning model.

Comparisons between (d-RSA + Priors + Planning) vs. (d-RSA + Priors) isolate (A), while (d-RSA + Priors + Planning) vs. (d-RSA + Planning) isolate (B).

6 RESULTS

6.1 When is Pragmatic Planning most useful?

To understand our full model’s strengths, we compare achieved rewards across all 800 trials against the three model variants: d-RSA, d-RSA + Priors, and d-RSA + Planning. In particular, to extract generalisations about scenarios where our full d-RSA + Priors + Planning model produces the highest gain over each baseline, we fit a linear mixed-effect regression predicting total reward as a function of model, scenario, and user properties, and their interactions, with random intercepts for uncontrolled variance. We apply this same regression approach to all metrics reported in the following sections.

Pairwise comparisons show the d-RSA + Priors + Planning model outperforms all baselines ($p < 0.01$), with increases associated with both planning ($\beta = 0.31, SE = 0.02, p < .001$) and belief-tracking ($\beta = 0.22, SE = 0.02, p < .001$). These elements are particularly valuable in combination (interaction $\beta = 0.05, SE = 0.02, p = .019$): planning benefits most when user-specific beliefs are available. Benefits are largest in the highest-difficulty scenarios with many critical properties or critical properties emerging closer in time ($p < 0.01$; see Figures 2a, 2b).

²Detailed descriptions in the supplementary material.

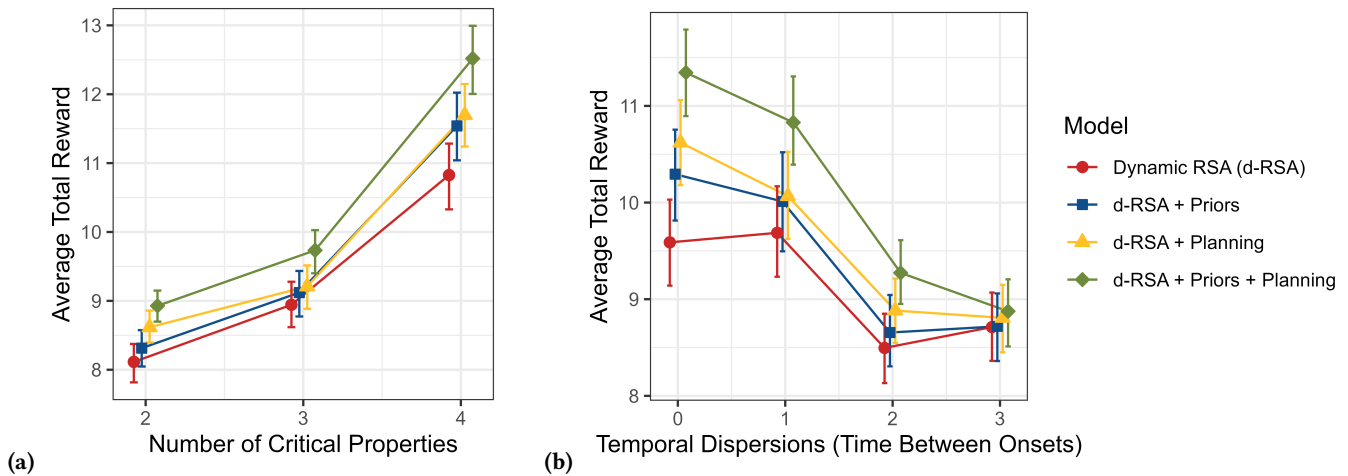


Figure 2: Total reward (y-axis) for each model variant across scenario factors. (a) Reward as a function of the number of critical properties (higher = more challenging). (b) Reward vs. temporal dispersion of critical onsets (lower = more tightly clustered, i.e., challenging). The full model achieves the greatest relative gains as scenarios become more difficult.

Unsurprisingly, we observe that modelling user beliefs delivers the highest gain when users begin with a higher awareness of critical properties ($p = 0.03$)—compare the scenario in Fig. 1, where user-based planning increases communicative speed by capitalising on partial pre-awareness. Table 1 presents cases where the model produces a higher value. In scenario (A), four properties become critical in short succession, beginning with two which the user anticipates. The simpler d-RSA and d-RSA + Planning models use long, specific utterances covering only two problems, failing to prioritise those that the user is unaware of. The d-RSA + Priors + Planning model packs in more messages by exploiting user knowledge for shorter alerts. Another strategy that emerges in crowded scenarios for the full model use of anticipatory alerts, raising awareness before criticality – in (B), where the d-RSA model waits silently before several properties become critical together, the d-RSA + Priors + Planning model uses specific messages directing attention to near-critical values. This still leaves space for a specific alert once criticality hits, while shifting the users’ priors for unanticipated problems, raising the chance that a quick *beep* is informative.

6.2 How is a message spoken?

We also investigate message choices by examining alert behaviour differences from the baseline. Message-type entropy diagnoses variability across specificity levels in each model’s highest-ranked sequence. The d-RSA + Priors + Planning model produces the highest entropy ($ps < 0.01$), particularly supported by planning ($p < 0.01$), while d-RSA and d-RSA + Priors models prefer fully-specific alerts ($ps < 0.01$). Belief-tracking increases entropy specifically in easier scenarios with fewer, more spread-out critical properties and for users who begin with better awareness ($ps < 0.01$).

Entropy is largely being driven by an increased adoption of less specific messages in the full model. Measuring message specificity as the ratio of unambiguous (two-feature) to ambiguous (single-feature alerts and beeps) messages, this ratio is lowest in the full

model ($ps < 0.01$), driven downward by both planning and belief-tracking, and especially by their combination ($ps < 0.01$). Both elements produce the largest reductions in specificity when users begin with good awareness ($ps < 0.01$; see Figure 3a). Planning furthermore reduces specificity when critical properties are temporally clustered ($p < 0.01$), when efficiency has the greatest value. See scenario (C), where the full d-RSA + Priors + Planning model uses less specific utterances for two better-known properties, reserving time to describe unexpected problems in detail. That is, when the model knows a user begins with priors that will lead them to resolve less specific messages correctly, it takes advantage of that knowledge, making it more likely to use non-specific messages.

In sum, our full model improves on less sophisticated systems by mixing complete descriptions with carefully-selected, less-specific alerts based on user knowledge and message density, using non-specific alerts when they would be correctly interpreted, where possible.

6.3 When is a message spoken?

Finally, we investigate how the model times the messages it chooses to send, using the median delay between the onset of low-awareness critical properties and the model’s first message expected to shift the distribution of user beliefs towards that property’s true value, for each trial.³ The d-RSA + Priors + Planning model achieves shorter delays than all others ($ps < 0.01$). Both planning and belief-tracking drive delay down ($ps < 0.01$), especially in combination ($p = 0.01$), with belief-tracking bringing the d-RSA + Priors model to second place ($ps < 0.01$), outranking the d-RSA + Planning and d-RSA models ($ps < 0.01$). Stronger belief-tracking improvements occur in particularly difficult cases with many critical properties ($p = 0.01$; see Figure 3b), dense temporal packing ($p < 0.01$), and higher initial user awareness ($p < 0.01$; see scenario (C).

³Critical properties never mentioned before the horizon are counted as maximum delay of 7.

Table 1: Example test scenarios, in terms of critical property identities and onsets, and the sequences generated by our models.

Critical Properties with Onset Time (User Unaware)	Top Sequence [U_B^* = Baseline (d-RSA), U_F^* = Full (d-RSA++)]	$\sum r_t \uparrow$
(A) (D2_Dist, 2), (D3_Alt, 3), (D3_Dist, 4), (D4_NoFly, 5)	$U_B^* = [\dots, 'D2 Distance', (X), (X), 'D3 Altitude', (X), (X)]$ $U_F^* = [\dots, 'Beep', 'Altitude', (X), 'D3 Distance', (X), (X)]$	9.65 12.29
(B) (D1_Wind, 3), (D1_Dist, 3), (D3_NoFly, 3), (D4_Alt, 3)	$U_B^* = [\dots, \dots, 'D1 Wind Speed', (X), (X), 'Distance', (X)]$ $U_F^* = ['D4 Altitude', (X), (X), 'D3 No Fly Zone', (X), (X), 'Beep']$	5.29 11.24
(C) (D1_Batt, 1), (D2_Rotor, 2), (D2_Alt, 3), (D2_NoFly, 4)	$U_B^* = ['D1 Battery', (X), (X), 'D2 Altitude', (X), (X), 'Beep']$ $U_F^* = ['Battery', (X), 'D2 Rotor', (X), (X), 'Altitude', (X)]$	13.74 19.08
(D) (D1_Batt, 1), (D1_Wind, 1), (D3_Alt, 1)	$U_B^* = ['D1 Battery', (X), (X), 'D1 Windspeed', (X), (X), 'Beep']$ $U_F^* = ['D3 Altitude', (X), (X), 'D1 Battery', (X), (X), \dots]$	8.98 13.47

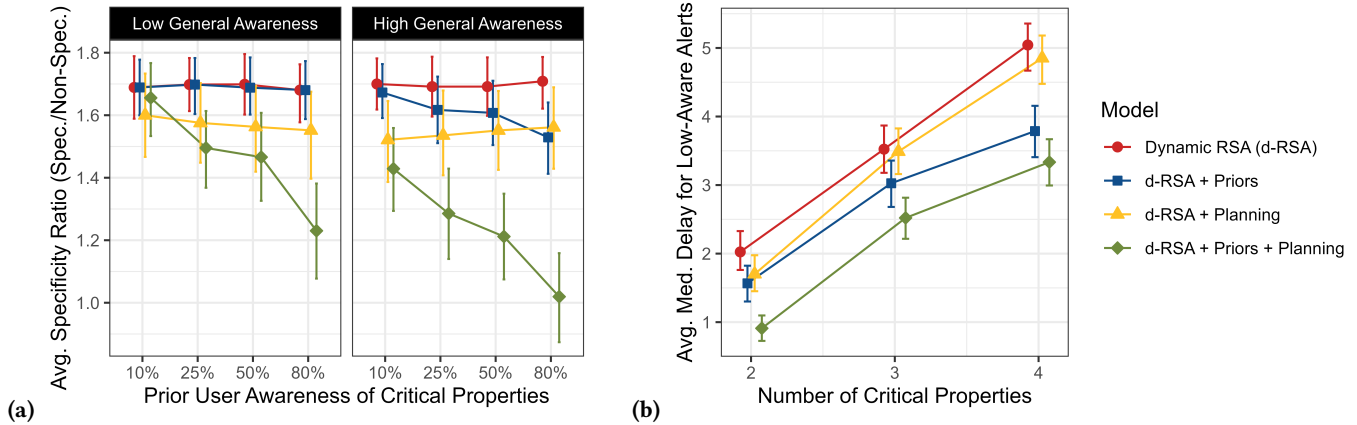


Figure 3: (a) Message Specificity Ratio (y-axis) vs. Prior User Awareness of critical properties (x-axis). Specificity decreases with higher prior awareness only in the full model, showing its adaptive use of less specific alerts; baseline variants remain less flexible. (b) Median Delay (in timesteps, y-axis) to the First Alert about Low-Awareness Critical Properties vs. Number of Critical Properties (x-axis). The full model is consistently faster as scenario difficulty grows, outperforming baseline variants.

Substantial delays indicate some risk of delayed response to critical situations. To check how this risk is being managed, we compare this against the median delay for high-awareness critical properties, taking their difference as a proportion of total delay across both (low vs. high awareness) categories. As discussed at the beginning (Figure 1), an adaptive agent should sometimes risk delays for likely-conscious problems to prioritise unconscious ones. Only our full d-RSA + Priors + Planning model significantly prioritises low-awareness properties (95% CI : 0.02, 0.13), as per the difference metric. Belief-tracking is the major contributor ($p < 0.01$), with the d-RSA + Priors model performing well but not significantly favouring low-awareness properties (95% CI : -0.04, 0.06) nor differing from the d-RSA + Priors + Planning model ($p = 0.12$). See scenario (D), where user prior awareness allows our d-RSA + Priors + Planning model to accurately prioritise among three simultaneous problems.

To summarise, our full model improves the timing of the most important alerts by selectively delaying mention of some problems, often prioritising those which the user is not aware of.

7 DISCUSSION AND CONCLUSION

We presented a pragmatic signalling framework that extends Rational Speech Acts (RSA) with temporal belief tracking, user-prior modelling, and finite-horizon planning for dynamic human-AI collaboration. Our systematic evaluation demonstrates that pragmatic planning – incorporating both user-awareness and lookahead reasoning – significantly improves communication effectiveness over baseline approaches, particularly in scenarios with challenging environments, diverse user profiles and temporal constraints. This work establishes a foundation for adaptive assistive systems that strategically time and frame alerts to align with individual users’ evolving situational awareness, a capability necessary for safety-critical domains where communication breakdowns have severe consequences. Supplementary Material and Appendices are available at <https://arxiv.org/abs/2510.23340>.

7.1 Our Contributions

Our work makes four key theoretical and methodological contributions to the broader landscape of human-AI collaboration research. First, we **extend the RSA framework to dynamic settings**. Classical RSA treats utterances as isolated communicative acts, ignoring

that temporal dynamics are essential to effective human-AI collaboration [11]. Our d-RSA framework represents the first systematic extension of the pragmatic reasoning capabilities of vanilla RSA across multiple timesteps. By modelling how listener beliefs propagate, d-RSA enables principled reasoning about communication sequences rather than individual messages.

Second, to the best of our knowledge, our work is the first to **jointly optimise** the *what*, *how*, and *when* of assistive communication while being **mindful of human cognition**. Most existing approaches in human-AI collaboration optimise either message content (framing) or temporal allocation (scheduling) independently. Information-theoretic methods focus on optimal content selection given fixed timing constraints [16, 23], while scheduling systems optimise temporal delivery assuming predetermined message content [5, 37]. Recent work points out their interdependence [30]: message framing directly impacts user comprehension, while timing determines the cognitive window available for processing.

We bridge this gap by recognising communication as a sequential planning problem where every choice has a temporal opportunity cost (see Eq. 8). Choosing to deliver a detailed, specific utterance may preclude timely alerts about other emerging problems, a trade-off that isolated optimisers cannot capture. By **jointly reasoning about content, framing, and timing**, our d-RSA framework enables the AI agent to act as a **cognitive interface**, guiding human attention — a finite resource — to where it is most needed in dynamic, high-stakes environments.

Third, our systematic evaluation of 800 diverse simulated trials helps reveal when pragmatic human-aware planning is most beneficial. Our results show it is most advantageous in precisely the scenarios where **adaptability** matters most: **high-complexity environments** with multiple simultaneous critical properties, diverse user knowledge states, and tight time constraints. This communication strategy exhibits **context-appropriate** behaviour consistent with Gricean cooperative principles [9] — using concise alerts when users possess relevant background knowledge (Maxim of Quantity), providing detailed explanations when users lack context, and strategically timing communications to maximise comprehension (Maxim of Manner). Furthermore, our analysis of anticipatory alerting indicates sophisticated reasoning beyond mere reactivity: the system proactively raises awareness of near-critical conditions, positioning users to interpret subsequent brief alerts more effectively; thereby reducing alert frequency and presumably the resulting alarm fatigue [19]. This **adaptive ability** to plan communication sequences hints at a cooperative partnership currently missing from many human-AI systems.

Fourth, our work provides a **principled theoretical grounding for assistive agents**. We contend that effective human-AI systems are built upon explicit, structured models of the task, the world, and the human partner. This ensures that cooperative behaviour is by design, and not a brittle, emergent property common to agents that rely on heuristic or purely data-driven methods [8, 21]. A growing body of research emphasises aforementioned Gricean norms provide an essential theoretical foundation for collaboration, with norm-adherent agents achieving higher task accuracy and generating more appropriate responses [38, 45]. Our choice to build upon the RSA framework stems directly from this insight. As a computational model of Gricean cooperation, RSA’s Bayesian recursive

structure provides the mechanism for Theory of Mind reasoning, allowing agents to explicitly model the mental states of their communicative partners [11].

This principled grounding offers a crucial advantage in reliability and predictability, particularly when contrasted with the often opaque reasoning of purely neural models and LLMs. While LLMs excel at generating fluent referring expressions [29], they also exhibit systematic pragmatic violations, such as excessive specificity, and can produce user misunderstandings more frequently than human partners [22, 42]. Our approach ensures adherence to cooperative principles, vital for safety-critical applications. The primary trade-off, of course, is the need for hand-crafted lexicons, a limitation that itself opens up several promising directions for future research.

7.2 Future Directions

Most immediately, this points toward hybrid approaches — combining RSA’s principled reasoning with the generative flexibility of neural models, reducing manual lexicon construction while maintaining pragmatic consistency [6]. Next, future work should relax our model’s assumptions of perfect knowledge of user awareness and future world states — idealisations that enable controlled comparison in simulated environments but may not hold in practice. On the user side, our static knowledge model can be extended by incorporating dynamic user attention models, enabling real-time adaptation through behavioural monitoring (e.g., eye-tracking [10]). On the environment side, moving beyond the assumption of a perfectly known future requires addressing the fundamental challenges of planning under uncertainty. Monte Carlo tree search [24, 26] approximations could enable scalable deployment while managing the uncertainty inherent in real-world environments. Theoretically, extending our binary criticality model to multi-level urgency hierarchies would better reflect real-world operational priorities — distinguishing, for e.g., between a gradual battery depletion and a sudden rotor malfunction.

To conclude, by integrating principled pragmatic reasoning with temporal planning, this work offers a path toward assistive agents that truly cooperate with human cognition rather than overwhelming it. In an era where human-AI interdependence gradually becomes inescapable, ensuring our AI agents are effective and collaborative is a scientific necessity. Our theoretically grounded approach aims to provide a concrete step in that direction.

ACKNOWLEDGMENTS

This work was supported by German Research Foundation, DFG grant 389792660, as part of the TRR 248 CPEC — Center for Perspicuous Computing, see <https://perspicuous-computing.science>.

REFERENCES

- [1] Jessica S Ancker, Alison Edwards, Sarah Nosal, Diane Hauser, Elizabeth Mauer, Rainu Kaushal, and With the HITEC Investigators. 2017. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC medical informatics and decision making* 17, 1 (2017), 36.
- [2] Carolyn Jane Anderson. 2021. Tell Me Everything You Know: A Conversation Update System for the Rational Speech Acts Framework. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2021*. 244–253.
- [3] Jacob Andreas and Dan Klein. 2016. Reasoning about pragmatics with neural listeners and speakers. *arXiv preprint arXiv:1604.00562* (2016).

- [4] Immanuel Barshi. 2023. People are the Weak Link in the System. (2023). Invited Talk at Ben Gurion University.
- [5] Mohan Baruwal Chhetri, Shahroz Tariq, Ronal Singh, Fatemeh Jalalvand, Cecile Paris, and Surya Nepal. 2024. Towards human-ai teaming to mitigate alert fatigue in security operations centres. *ACM Transactions on Internet Technology* 24, 3 (2024), 1–22.
- [6] Reuben Cohn-Gordon, Noah D. Goodman, and Christopher Potts. 2017. Pragmatically Informative Image Captioning with Character-Level Inference. In *Proceedings of NAACL-HLT 2018*. 439–443. <https://doi.org/10.18653/v1/N18-2070>
- [7] Reuben Cohn-Gordon, Noah D. Goodman, and Christopher Potts. 2019. An Incremental Iterated Response Model of Pragmatics. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2019*. 81–90.
- [8] Katherine M. Collins, Iliia Sucholutsky, Umang Bhatt, Kartik Chandra, Lionel Wong, Mina Lee, Cedegao E. Zhang, Tan Zhi-Xuan, Mark Ho, Vikash Mansinghka, Adrian Weller, Joshua B. Tenenbaum, and Thomas L. Griffiths. 2024. Building Machines that Learn and Think with People. *Nature Human Behavior* 8, 10 (2024), 1851–1863. <https://doi.org/10.1038/s41562-024-01991-9>
- [9] Robert Dale and Ehud Reiter. 1995. Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive science* 19, 2 (1995), 233–263.
- [10] Anwsha Das, Zekun Wu, Iza Skrjanec, and Anna Maria Feit. 2024. Shifting Focus with HCEye: Exploring the Dynamics of Visual Highlighting and Cognitive Load on User Attention and Saliency Prediction. *Proceedings of the ACM on Human-Computer Interaction* 8, ETRA (2024), 1–18.
- [11] Judith Degen. 2023. The Rational Speech Act Framework. *Annual Review of Linguistics* 9 (2023), 519–540. <https://doi.org/10.1146/annurev-linguistics-031220-010811>
- [12] Mica R Endsley. 2017. Toward a theory of situation awareness in dynamic systems. In *Situational awareness*. Routledge, 9–42.
- [13] Michael C. Frank and Noah D. Goodman. 2012. Predicting Pragmatic Reasoning in Language Games. *Science* 336 (2012), 998.
- [14] Michael Franke. 2011. Quantity Implicatures, Exhaustive Interpretation, and Rational Conversation. *Semantics & Pragmatics* 4 (2011), 1. <https://doi.org/10.3765/sp.4.1>
- [15] Daniel Fried, Jacob Andreas, and Dan Klein. 2017. Unified pragmatic models for generating and following instructions. *arXiv preprint arXiv:1711.04987* (2017).
- [16] Piotr J Gmytrasiewicz and Edmund H Durfee. 2001. Rational communication in multi-agent environments. *Autonomous Agents and Multi-Agent Systems* 4, 3 (2001), 233–272.
- [17] Noah D. Goodman and Michael C. Frank. 2016. Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences* 20, 11 (2016), 818–829. <https://doi.org/10.1016/j.tics.2016.08.005>
- [18] Noah D. Goodman and Andreas Stuhlmüller. 2013. Knowledge and Implicature: Modeling Language Understanding as Social Cognition. *Topics in Cognitive Science* 5 (2013), 173–184. <https://doi.org/10.1111/tops.12007>
- [19] Muaz Gülşen, Arslan, Sevban, and et al. 2025. The Effect of Alarm Fatigue on the Tendency to Make Medical Errors in Surgical Intensive Care Nurses: A Correlational Study Examining the Role of Moderating Factors. In *Healthcare*, Vol. 13. MDPI, 631.
- [20] Robert D. Hawkins, Hyowon Gweon, and Noah D. Goodman. 2021. The Division of Labor in Communication: Speakers Help Listeners Account for Asymmetries in Visual Perspective. *Cognitive Science* 45 (2021), e12926. <https://doi.org/10.1111/cogs.12926>
- [21] Prakash Jamakatel, Rebecca De Venezia, Christian Muise, and Jane Jean Kiam. 2024. A goal-directed dialogue system for assistance in safety-critical application. *International Joint Conference on Artificial Intelligence* (2024).
- [22] Rhea Kapur, Robert Hawkins, D., and Elisa Kreiss. 2025. When More Words Say Less: Decoupling Length and Informativity in Image Description Evaluation. (2025). Unpublished manuscript, Stanford.
- [23] Seth Karten, Siva Kailas, and Katia P Sycara. 2023. Emergent Compositional Concept Communication through Mutual Information in Multi-Agent Teams.. In *AAMAS*. 2391–2393.
- [24] Thomas Keller and Malte Helmert. 2003. Trial-Based Heuristic Tree Search for Finite Horizon MDPs. In *Proceedings of the 13th International Conference on Automated Planning and Scheduling (ICAPS'03)*, Enrico Giunchiglia, Nicola Muscettola, and Dana Nau (Eds.).
- [25] Md. Nafiz Hasan Khan and Carman Neustaedter. 2019. An Exploratory Study of the Use of Drones for Assisting Firefighters During Emergency Situations. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 1–14. <https://doi.org/10.1145/3290605.3300502>
- [26] Levente Kocsis and Csaba Szepesvári. 2006. Bandit Based Monte-Carlo Planning. In *Proceedings of the 17th European Conference on Machine Learning (ECML 2006)*, Johannes Fürnkranz, Tobias Scheffer, and Myra Spiliopoulou (Eds.), Vol. 4212. Springer-Verlag, 282–293.
- [27] Jessy Lin, Nicholas Tomlin, Jacob Andreas, and Jason Eisner. 2024. Decision-Oriented Dialogue for Human-AI Collaboration. *Transactions of the Association for Computational Linguistics* 12 (Aug. 2024), 892–911. https://doi.org/10.1162/tacl_a_00679
- [28] Dennis Victor Lindley. 1972. *Bayesian statistics: A review*. SIAM.
- [29] Haotian Liu, Chunyuan Li, Yuheng Li, and Yong Jae Lee. 2024. Improved Baselines with Visual Instruction Tuning. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 26286–26296. <https://doi.org/10.1109/CVPR52733.2024.02484>
- [30] Shipeng Liu, FNU Shrutika, Boshen Zhang, Zhehui Huang, and Feifei Qian. 2024. Effect of adaptive communication support on human-ai collaboration. AAAI 2025 Workshop on Advancing LLM-Based Multi-Agent Collaboration.
- [31] Ziqiao Ma, Jing Ding, Xuejun Zhang, Dezhi Luo, Jiahe Ding, Sihan Xu, Yuchen Huang, Run Peng, and Joyce Chai. 2025. Vision-Language Models Are Not Pragmatically Competent in Referring Expression Generation. (2025). <https://doi.org/10.48550/arXiv.2504.16060> arXiv manuscript.
- [32] Erik Miehling, Manish Nagireddy, Prasanna Sattigeri, Elizabeth M. Daly, David Piorkowski, and John T. Richards. 2024. Language Models in Dialogue: Conversational Maxims for Human-AI Interactions. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen (Eds.). Association for Computational Linguistics, Miami, Florida, USA, 14420–14437. <https://doi.org/10.18653/v1/2024.findings-emnlp.843>
- [33] John Mongan and Marc Kohli. 2020. Artificial intelligence and human life: five lessons for radiology from the 737 MAX disasters. , e190111 pages.
- [34] Will Monroe, Robert D. Hawkins, Noah D. Goodman, and Christopher Potts. 2017. Colors in Context: A Pragmatic Neural Model for Grounded Language Understanding. *Transactions of the Association for Computational Linguistics* 5 (2017), 323–338. https://doi.org/10.1162/tacl_a_00064
- [35] Will Monroe, Robert XD Hawkins, Noah D Goodman, and Christopher Potts. 2017. Colors in context: A pragmatic neural model for grounded language understanding. *Transactions of the Association for Computational Linguistics* 5 (2017), 325–338.
- [36] Marco Mussi, Alberto Maria Metelli, Marcello Restelli, Gianvito Losapio, Ricardo J Bessa, Daniel Boos, Clark Borst, Giulia Leto, Alberto Castagna, Ricardo Chavarriaga, et al. 2025. Human-AI interaction in safety-critical network infrastructures. *iScience* 28, 9 (2025).
- [37] Stephanie Rosenthal and Laura M Hiatt. 2020. Human-centered decision support for agenda scheduling. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*. 1161–1168.
- [38] Fardin Saad, Pradeep K Murukannaiah, and Munindar P Singh. 2025. Gricean Norms as a Basis for Effective Collaboration. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*. 1812–1820.
- [39] Gregory Scontras, Michael Henry Tessler, and Michael Franke. 2021. A practical introduction to the Rational Speech Act modeling framework. *arXiv preprint arXiv:2105.09867* (2021).
- [40] Scott Shappell, Cristy Detwiler, Kali Holcomb, Carla Hackworth, Albert Boquet, and Douglas Wiegmann. 2006. Human Error and Commercial Aviation Accidents: A Comprehensive, Fine-Grained Analysis Using HFACS. (2006).
- [41] Cyrus Sigari and Peter Biberthaler. 2021. Medical Drones: Disruptive Technology Makes the Future Happen. *Unfallchirurg* 124, 12 (2021), 974–976. <https://doi.org/10.1007/s00113-021-01095-3>
- [42] Zineng Tang, Lingjun Mao, and Alane Suhr. 2024. Grounding Language in Multi-Perspective Referential Communication. In *Proceedings of EMNLP 2024*. 19727–19741. <https://doi.org/10.18653/v1/2024.emnlp-main.1100>
- [43] Christopher D Wickens. 2008. Multiple resources and mental workload. *Human factors* 50, 3 (2008), 449–455.
- [44] Arianna Yuan, Will Monroe, Yue Bai, and Nate Kushman. 2018. Understanding the Rational Speech Act Model. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 40.
- [45] Tan Zhi-Xuan, Lance Ying, Vikash Mansinghka, and Joshua B Tenenbaum. 2024. Pragmatic Instruction Following and Goal Assistance via Cooperative Language-Guided Inverse Planning. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*. 2094–2103.