An ACT-R model of resource-rational performance in a pragmatic reference game

John Duff[1], **Alexandra Mayn**[1] and **Vera Demberg**[1,2]
{amayn, jduff, vera}@lst.uni-saarland.de
[1] Department of Language Science and Technology, Saarland University
[2] Department of Computer Science, Saarland University

Abstract

In the Gricean tradition, pragmatic competence is part of the general human capacity for social reasoning. Indeed, human performance in reference games involving ad-hoc implicatures sometimes aligns with idealized models of rational interaction. But such experiments have also found that humans derive far fewer implicatures than ideal models, subject to individual differences unrelated to social reasoning. In this paper, we consider whether these patterns could arise from the resource-rational deployment of a core social competence, such that individuals choose from various strategies of interpretation, given those strategies' resource demands and success rates, subject to individually-varying predispositions and exploration tendencies. We construct a model of this resource-rational performance in the cognitive architecture ACT-R—to our knowledge the first mechanistic model of performance in these tasks—and we examine its predictions for multi-trial reference games across two model experiments. The model reproduces the key patterns in the human data, providing an initial proof of concept for the role of resource-rationality in these tasks and opening a new avenue for understanding individual differences in pragmatic reasoning.

Keywords: pragmatic reasoning; resource-rationality; ACT-R

Introduction

In many conversations, we readily interpret and act on utterances which are literally under-informative. For instance, if a busy baker has loaves of bread with rosemary, with olives and walnuts, and with olives alone, they are likely to hand you the latter if you ask for "bread with olives." They would be making a good guess to your intentions; intuitively, if you wanted the option with walnuts too, you would have said something else. Per Grice (1975), and much subsequent work, many inferences can be thought of in this way, as part of a general capacity for social inference, considering the expected behavior of cooperative, goal-oriented agents.

Recently, this approach has been evaluated by adopting formalizations of social inference using the tools of game theory and Bayesian reasoning (Degen, 2023; Franke, 2011; Goodman & Frank, 2016; Jäger, 2010; Parikh, 1991; Van Rooy, 2004) and considering how well such models predict production and comprehension behavior in simple "reference games" (Carstensen et al., 2014; Degen & Franke, 2012; M. C. Frank & Goodman, 2012; M. C. Frank et al., 2016; Qing & Franke, 2015; Rohde et al., 2012; Stiller et al., 2011, 2015). When the data match the predictions of these models, they seem to support the core Gricean hypothesis linking communicative inference with general social reasoning.

But what should we conclude when an experiment reveals behavior that these idealized accounts cannot explain? In fact, participants in single-trial reference games systematically neglect certain predicted inferences (Sikos et al., 2021), which only seem to ever emerge as a factor of experience in multi-trial games (Degen & Franke, 2012)—and which even then are subject to substantial individual differences, including variance linked to non-social problem-solving tasks (Mayn & Demberg, 2023b). These patterns seem to be at odds with the idea of a robust capacity for social reasoning.

In this paper, we emphasize that in order to compare observed behavior to theories of capacity, we need also a non-trivial theory of performance, the goals and limitations of the agent who is exercising that capacity. In particular, we consider the hypothesis that comprehenders in an actual interaction engage in *resource-rational adaptation* (Hawkins et al., 2021; Howes et al., 2009; Lieder & Griffiths, 2020): they explore a space of possible interpretation procedures, driven by efficiency. We present a model of this type of comprehender in the ACT-R cognitive architecture (Anderson & Lebiere, 1998), and explore its behavior under a variety of parameter settings, comparing against the patterns that we aim to explain from prior literature. In Experiment 1, we show that the model correctly predicts that certain implicatures should be initially unlikely, but increase over experience in the task. Crucially, the model also can generate patterns of individual differences which would correlate with problem-solving performance, by varying parameters controlling the dynamics of strategy exploration (Stocco et al., 2021). In Experiment 2, we also vary agents' predisposition towards pragmatic strategies, as a proxy for variance in Theory of Mind, producing not only an expected correlation, but also an interesting sub-additive interaction with problem-solving. On the whole, we see our results as a promising proof of concept for how performance demands can explain the complexities of the behavior we see in reference game studies.

Background

Previous experiments using the reference game paradigm have revealed four key generalizations that we consider crucial for a successful model of behavior in these games: (a) asymmetries of difficulty among trial types, (b) task adaptation, (c) individual differences linked to Theory of Mind, and (d) individual differences linked to problem-solving.

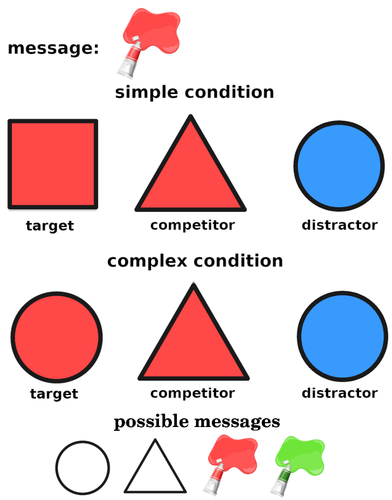

Figure 1: Examples of Simple and Complex trials in this type of reference game.

Task. The tasks we will discuss and model mainly involve single-message iconic communication as in Figure 1, with comprehenders asked to interpret a message sent by an (imaginary) interlocutor attempting to indicate one of three possible referents in the common ground (M. C. Frank & Goodman, 2012). In critical trials, the literal value of the message matches two referents, the *target* and *competitor*, but an inference is available which could lead to a preference to select the *target*. In this particular version of the game, the message comes from a constrained bank (also common ground); some features cannot be communicated (here, 'blue' and 'square').

Degen and Franke (2012) introduce this version of the game in order to compare performance across two types of trials. In Simple trials, the message bank allows only one matching message for the target (the observed message), but two for the competitor. In this case, even reasoning about the expected behavior of a purely literal speaker ('first-order pragmatic interpretation') will reveal a difference in likelihood in favor of the target—e.g. in Fig. 1, 'red' messages will be more likely from a speaker describing the red square (who will always send this message) vs. one describing the red triangle (who should only send 'red' half the time). In Complex trials, the target and competitor both match two messages, but an inference can still be drawn if one considers the expected behavior of a cooperative speaker ('second-order pragmatic interpretation,' the type of reasoning assumed in Grice, 1975 and M. C. Frank and Goodman, 2012)—in Fig. 1, because the competitor has an alternative unambiguous message ('triangle'), 'red' messages will be infrequent for that competitor and relatively more likely for the target.

Asymmetries and Adaptation. Across many studies, Simple trials are seen to elicit many more implicatures than Complex (Degen & Franke, 2012; M. C. Frank et al., 2016; Franke & Degen, 2016; Stiller et al., 2011); Franke and Degen (2016) report a drop from 77% target selections in Simple trials to 57% in Complex. They take this asymmetry as evidence that second-order pragmatic interpretation is rarely applied.

In particular, evidence suggests that second-order pragmatic interpretation is only ever applied when participants can realize through repeated exposure that it is necessary. Target selection on Complex trials is even less frequent in single- or few-trial experiments, where several studies have found no evidence of second-order interpretation (M. C. Frank et al., 2016; Sikos et al., 2021; Stiller et al., 2011, though cf. partial evidence in Qing and Franke, 2015). Indeed, several longer studies report evidence that performance increases across both types of critical trial over time (Degen & Franke, 2012; Mayn & Demberg, 2023a, 2023b).

Individual Differences. Within these general trends, research has also revealed that comprehenders are remarkably diverse in their individual game performance. While Franke and Degen (2016) observe that participants filling the role of the speaker are well-categorized as following a single strategy, participants playing as comprehenders varied widely, with evidence for at least three classes corresponding to expected performance for first-order interpretation (succeeding on Simple but at chance on Complex), second-order interpretation (succeeding at both, somewhat rare), and even purely literal interpretation (at chance on both, not uncommon). A recent follow-up by Mayn and Demberg (2023b; see also earlier results in Mayn and Demberg, 2022) explored potential sources for this variation, subjecting participants to a battery of several additional tasks intended to measure individual differences in working memory, Theory of Mind, and non-social problem-solving. Figure 2 illustrates the two patterns they observe: (1) a main effect of Theory of Mind ability (ToM), as measured by the Reading the Mind in the Eyes task (Baron-Cohen et al., 2001) and the Short Story task (Dodell-Feder et al., 2013), associated with higher rates of target selections across both conditions, and also (2) effects of problem-solving ability, as measured by a sample of Raven's Progressive Matrices (Raven et al., 1998) and the Cognitive Reflection Test (Frederick, 2005), associated with higher rates of target selections, most strongly in the Simple condition.

The former effect is unsurprising, under a Gricean account where these inferences are supported by reasoning about the intentions of other agents, and also given previous findings of relationships between ToM task performance and other pragmatic tasks (Fairchild & Papafragou, 2021; Trott & Bergen, 2019). But the latter requires some explanation, as the kinds of careful reasoning required by these problem-solving tasks do not readily map onto any component of our theories for pragmatic competence.

This connection can be better understood by examining the literature in cognitive psychology on Raven's Matrices, which has identified general constructs that may underlie variance in the task at a finer level than the notion of an abstract problem-solving ability. One source of individual differences in variation is the identification of a suitable strategy for completing the task (Gonthier & Thomassin, 2015;

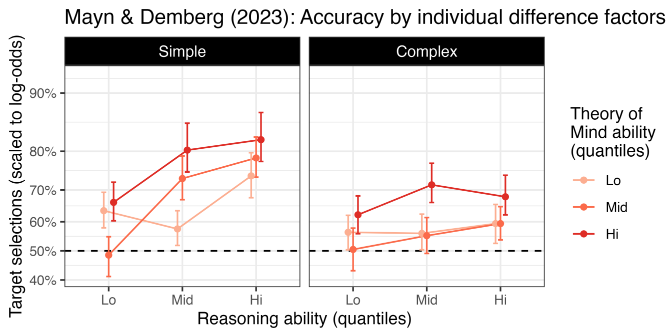

Figure 2: Relationships in Mayn and Demberg (2023b) between target selection and PCA constructs for Theory of Mind ability and problem-solving ability. In all figures, error bars indicate bootstrapped 95% CIs, and the dashed line marks chance (guessing between target and competitor).

Gonthier et al., 2024; Hayes et al., 2011; Jarosz & Wiley, 2012; Vigneau et al., 2006). Gonthier et al. (2024) report that in their sample, more than half of developmental improvement in Raven's Matrices performance with age can be explained by better strategy selection. Another source, highlighted in recent work by Stocco et al. (2021), is how participants explore the space of possible solutions to each problem, both their *persistence*—how long they are willing to cycle through possible hypotheses before giving up (see also Cheyette and Piantadosi, 2024)—and how quickly they disengage from failed hypotheses (cf. Storm et al., 2011). Stocco et al. (2021) formalize the latter as the *strength of negative feedback* (F_{NEG}) self-applied within a reinforcement learning system (M. J. Frank et al., 2004), which they implement in ACT-R, and validate with a secondary task. We adopt the idea here that much variation in complex problem-solving tasks is determined by domain-general resources for exploration, and in particular, persistence and F_{NEG}.

Modeling in ACT-R

We propose that we can account for all four key patterns (asymmetries of difficulty, task adaptation, and individual differences linked to both Theory of Mind and general problem-solving) as consequences of resource-rational adaptation during performance of this task. In doing so, we follow a common modern position on the basic goal-oriented flexibility of cognition (Howes et al., 2009; Lieder & Griffiths, 2020): that individuals seek through task experience to maximize the expected utility of their strategy for the task, where a strategy's utility is crucially a function of both the likelihood of success, and the resources and time required to carry out the strategy. Hawkins et al. (2021) use resource-rationality to explain very similar patterns in a viewpoint asymmetry task, where comprehenders initially neglect the speaker's perspective, and subsequently adapt. Applying this story here, we consider that comprehenders begin from a state of efficiency-motivated pragmatic reluctance, but subsequently lower expectations about the utility of less pragmatic strategies, identify alternative strategies to try, and raise expectations about the utility of those. This process should lead to changes in behavior over time, with starting points dependent on initial predispositions (which we take to be related to performance on ToM tasks), and degrees of improvement dependent on the dynamics of exploration behavior (which we take to be related to performance in problem-solving tasks).

We examine the quantitative predictions of this account by creating a model of a resource-rational comprehender in a multi-trial reference game task. To test the account as stated above, this model has two necessary components: (1) it should implement several strategies for the task realistically given the constraints of real-time human cognition, so that they can be meaningfully compared in terms of their realistic execution time, and (2) it should implement some individually-parameterized mechanism for exploring those strategies over task performance and updating expected utility accordingly. We choose to specify our model in ACT-R (Anderson & Lebiere, 1998) because it is well-suited to these goals: it is designed to account for not only the outcome but also the plausible timecourse of various processes, and it implements a parameterizable mechanism for utility learning and exploration.[1] The goal of efficiency is built into the latter by the use of temporal difference learning (Fu & Anderson, 2006), penalizing strategies which only lead to success at a long delay. The model itself was implemented in Python using the `pyactr` library (Brasoveanu & Dotlačil, 2020).[2]

Design. ACT-R models treat cognitive processes like problem-solving as the firing of sequences of conditioned rules, which direct sensory, memory, and motor operations, shifting information between modules of the architecture in steps that consume various amounts of time. We equip the basic model with rules that produce three such sequences corresponding to the mechanical implementation of literal, first-order, and second-order interpretation in this task. The sequence corresponding to literal interpretation merely compares an observed message with each possible referent, taking about 1s to successfully interpret an unambiguous message. The first-order sequence also compares matching referents by the number of alternative messages available in the message bank, taking about 3.5s in a Simple trial. Finally, the second-order sequence further compares those alternative messages by the number of other referents they could match, taking about 4.0s in a Complex trial.

As schematized in Figure 3, the agent decides between these strategies by virtue of their utility, updated whenever they are attempted within a cycle of multiple attempts per trial. Each trial begins with an attempt at implementing the literal interpretation strategy. The agent, like humans in the

[1] By comparison, a non-mechanistic approach as in Hawkins et al. (2021) cannot compare resource demands in terms of execution time. Otherwise, individual differences in the task adaptation mechanism should be in principle possible in Bayesian-update models like Brochhagen (2021) or Hawkins et al. (2022), but we find it easier to explore these differences in the reinforcement learning idiom.

[2] See Duff et al. (2025) for our model code and simulation results.

experiments above, receives no signal of correct interpretation, but considers a strategy successful if it can identify a single preferred referent for the observed message, in which case it executes motor commands to select it as a response, and positive feedback triggers an update to that strategy's utility.[3] When strategies do not identify a preferred referent (e.g. when they cannot break a tie, as in the literal strategy as applied to the examples in Figure 1), utility is updated using negative feedback (F_{NEG})[4], and strategy selection repeats by selecting the current highest-utility strategy, subject to some noise.[5] If no successful strategy is found before the model's internal clock (Taatgen et al., 2007) exceeds its persistence parameter τ, this loop is interrupted by a forced guess.

Evaluation. We will evaluate this model here on simulated multi-trial experiments of the design introduced in Franke and Degen (2016), featuring 12 Simple and 12 Complex trials, plus 9 completely Ambiguous, and 33 completely Unambiguous fillers, randomly shuffled and with referents randomly ordered. Because the human data has only tracked final responses, our evaluation will focus on the model's response distributions. Nevertheless, we note that the model also generates predictions about gaze distribution and response times, which should be investigated in future work.

We report meaningful generalizations over the results by fitting Bayesian logistic regressions over target selection in critical trials using the R package `brms` (Bürkner, 2017), following the analysis used in Mayn and Demberg (2023b). Regressions included effects of condition (sum-coded, complex = +1.0), trial number (centered), scaled and centered individual difference parameters, and all of their two-way interactions. We take as notable any parameters whose 95% highest density posterior intervals exclude 0.

Model Experiment 1

We first investigate the predictions of the model while independently varying persistence and the strength of negative feedback (F_{NEG}), to examine whether factors that are hypothesized to control variance in problem-solving tasks like Raven's Matrices would also control variance in this task, given our model.

Method. Based on pilot simulations, we chose to examine performance with one of twenty F_{NEG} values between –0.05 and –10.00, and one of ten persistence values between 24 and 33, running 25 simulations for each combination of values for a total of 5000 simulations. Utility learning used the default learning rate of 0.2, and noise with the scale parameter 0.6. Initial utilities for literal, first-order, and second-order interpretation were 5.0, –2.5, and –5.0, respectively. For reference, in our model, with average F_{NEG}, a strategy with 100% success rate and no delay would approach a utility of 5.0, while strategies with 100% success rate and a 10s delay or a 50% success rate and a 5s delay would approach a utility of -5.0. Other free parameters influencing processing speed were set by common defaults and best practices for ACT-R modeling (ACT-R Research Group, 2022).

Results. The model reproduces the attested main effect of condition, with lower target selection rates in Complex trials than Simple trials, $\hat{\beta}$ = -1.00 [-1.03, -0.98].

As seen in Figure 4, the model also produces reliable positive relationships between target selection and both individual parameters for exploration: F_{NEG}, $\hat{\beta}$ = 0.76 [0.74, 0.78],[6] and persistence, $\hat{\beta}$ = 0.95 [0.92, 0.97]. Condition modulates both effects: comparable increases in F_{NEG} delivered smaller increases in target selection in Complex trials, $\hat{\beta}$ = -0.14 [-0.16, -0.13], as for persistence, $\hat{\beta}$ = -0.56 [-0.58, -0.54].[7]

Both conditions improve over task exposure, $\hat{\beta}$ = 0.03 [0.03, 0.03], equivalent to an increase of 2 logits over the course of the experiment, or an increase from 72% to 95% accuracy in critical trials. Additional interactions with F_{NEG}, $\hat{\beta}$ = 0.01 [0.01, 0.01], and persistence, $\hat{\beta}$ = 0.01 [0.01, 0.01], show that this adaptation effect was larger with parameters that allowed for faster adaptation (Figure 5).

Discussion. Our ACT-R model of resource-rational comprehension in a reference game is able to successfully reproduce conditional asymmetries and gradual improvements in target selection over multiple trials. In addition, by linking rates of adaptation to domain-general resources for rapid exploration of a search space, the model predicts a positive association between target selection in this task and accurate problem-solving in any other task which should require this type of exploration, as Stocco et al. (2021) argue for Raven's Matrices. This would seem to account for the association with problem-solving observed in Mayn and Demberg (2023b)—including the interaction with condition. The generality of the predicted association, and its dependence on these particular parameters, should be tested in future work.

A further prediction, rates of adaptation varying depending on exploration parameters, did not emerge as a significant interaction in Mayn and Demberg (2023b), although we note that an interaction as small as the one we observe here is likely difficult to detect under realistic conditions. More critically, we note that expected adaptation effects, and overall rates of target selection, are somewhat higher than those reported by Mayn and Demberg. This is likely due to an over-representation of highly-effective exploration strategies in our simulations here. We consider that actual participants

[3] We used a fixed F_{POS} of 5.0.

[4] The particular utility-learning algorithm as implemented in ACT-R leads to more extreme negative shifts for strategies which take longer to execute: $U_{n,t} = U_{n,t-1} + \alpha\Big((F-\delta) - U_{n,t-1}\Big)$, where $U_{n,t}$ is the utility of rule n at time t, α is the learning rate, F is the feedback received, and δ is the time elapsed since rule n last fired.

[5] When a failed strategy still has the highest utility after negative feedback has been applied, it will be resampled. We follow Stocco et al. (2021) in adding another loop of negative feedback in case of resampling, to simulate more effortful disengagement from strategies with this kind of persistent advantage.

[6] Here, higher values of F_{NEG} mean stronger negative feedback.

[7] The two parameters also show an unsurprising super-additive interaction: increases of F_{NEG} are increasingly effective at higher values of persistence, $\hat{\beta}$ = 0.26 [0.24, 0.28].

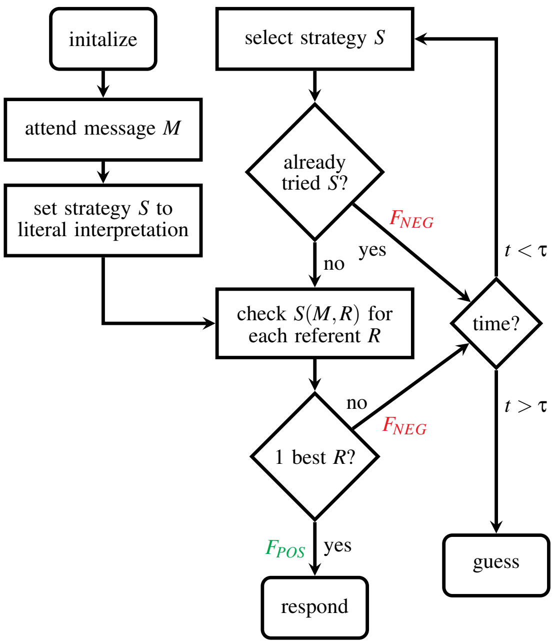

Figure 3: A schematic of our model's control flow.

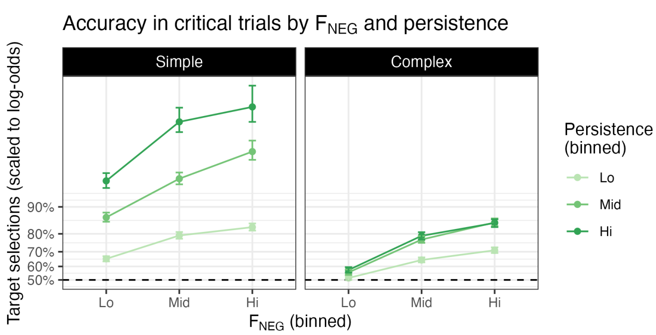

Figure 4: ACT-R model predictions for the relationship between target selection and individual differences of exploration (F_{NEG} and persistence, in bins of equal size and coverage), across critical conditions.

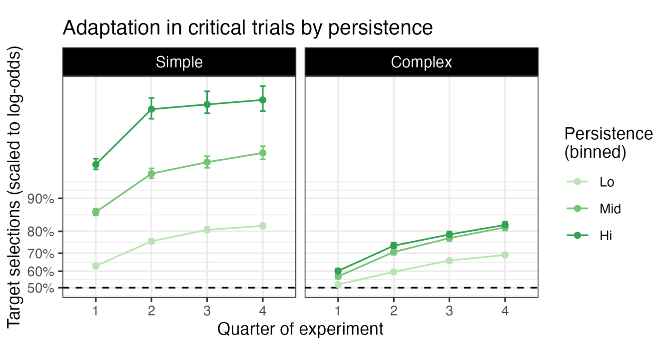

Figure 5: ACT-R model predictions for the relationship between target selection and task exposure (across quarters of the simulated experiment), mediated by persistence.

may not be uniformly distributed across the F_{NEG} and persistence values we considered here, but instead concentrated in the lower half or third of those ranges; we hope that future work can infer more about that distribution.

Model Experiment 2

We next investigate the predictions of the model while varying the initial expected utility of first- and second-order interpretation strategies, to determine whether this model can also capture hypothetical differences in predisposition towards pragmatic reasoning.

Method. Holding the starting utility of literal interpretation constant at 5.0, we vary the possible values for first- and second-order strategies (U_1 and U_2) within the plausible range of [0.0, -2.5, -5.0, -7.5, -10.0].[8] As we assume participants should always initially prefer first-order to second-order reasoning, we investigate only the ten combinations of values in that range where $U_1 > U_2$. We manipulated these utility settings in addition to the parameters above, with 5 simulations per unique combination, for a total of 10000 simulations.

Results. Regressions yielded similar relationships for all effects and interactions discussed for Experiment 1. In addition, we observe positive effects of U_1, $\hat{\beta} = 0.16$ [0.15, 0.18], and U_2, $\hat{\beta} = 0.36$ [0.34, 0.38]. Unsurprisingly, U_1 is particularly important for Simple trials, where it instantiates a minimal threshold for learning an effective strategy, $\hat{\beta} = -0.20$ [-0.21, -0.18], and likewise for U_2 in Complex trials, $\hat{\beta} = 0.05$ [0.03, 0.06]. It's hard to say exactly how one should map these separate utilities back to the global construct of ToM ability, where they are presumably closely correlated. We note that the marginal effect of U_1 on simple accuracy, $\hat{\delta} = 0.37$, is roughly equal to the marginal effect of U_2 on complex accuracy, $\hat{\delta} = 0.41$.

Moreover, these effects of initial estimated utility interact variously with the other two parameters. Figure 6 shows interactions between F_{NEG} with U_1, $\hat{\beta} = -0.09$ [-0.10, -0.08], and with U_2, $\hat{\beta} = 0.15$ [0.13, 0.16]. These are largely driven by boundary effects: notice that in the first panel, as U_1 increases, Simple accuracy quickly approaches a ceiling, reducing the effectiveness of F_{NEG}, while in the second panel, as U_2 increases, Complex accuracy moves away from a floor where F_{NEG} hardly mattered towards cases where learning can be much more effective.[9]

Discussion. Our model is able to jointly capture individual differences of exploration abilities, in parameters governing the dynamics of adaptation, with differences in predispositions to social reasoning, modeled here as differences in initial expected utilities for pragmatic interpretation. Beyond that basic proof of concept, we also see that the model expects a certain dependency between these factors: if participants have an extreme predisposition towards or away from

[8]To compare to the explanations of values above: a utility of 0.0 is equivalent to the belief that a strategy is e.g. 80% successful and takes 3s, while a utility of -10.0 maps to e.g. 50% success in 10s.

[9]Note that the regressions we fit cannot fully capture these nonmonotonic interactions, although the patterns are clear.

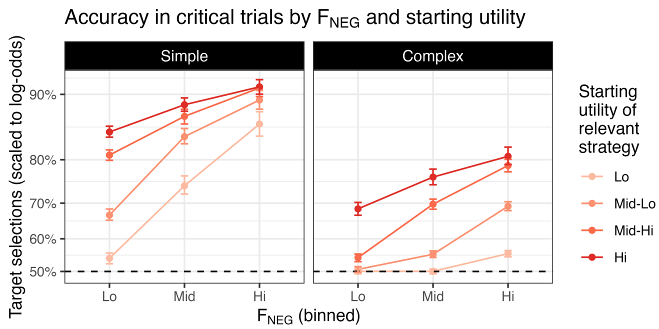

Figure 6: ACT-R model predictions for the relationship between target selection, F_{NEG}, and utilities (L: U_1, R: U_2).

pragmatic reasoning, the efficiency of their exploration has relatively little effect on their behavior. As a result, we would expect that the relationship between performance on this task and performance on general problem-solving tasks should be somewhat disrupted for individuals with especially low or high performance in ToM tasks. To examine whether this may be true, we perform a reanalysis of the Mayn and Demberg (2023b) data including the potential for such an interaction. We observe a small trend towards sub-additivity, $\hat{\beta}$ = –0.06 [–0.24, 0.11], such that a robust main effect of problem-solving, $\hat{\beta}$ = 0.34 [0.13, 0.36], is no longer credible for participants with ToM performance two standard deviations above the sample mean, $\hat{\beta}$ = 0.22 [–0.14, 0.58]. This is somewhat apparent in Figure 2, where slopes associated with problem-solving ability are less pronounced in the darker lines corresponding to the highest ToM ability, a pattern visible more clearly in our model's output (Figure 6). The Mayn and Demberg data thus do not rule out such an interaction, nor do they represent strong evidence in favor of one. We think that in the end, proper interpretation of these patterns requires a clearer vision for how these utility values should be related to global ToM performance. Nevertheless, we are interested to note that the model was able to generate a novel prediction for a complex dependency that may be empirically plausible.

General Discussion

We have constructed a mechanistic model of comprehension in a multi-trial reference game—to our knowledge, the first for any similar task[10]—allowing us to incorporate a theory of ideal Gricean competence within an explicit model of resource-rational adaptation. The results of our modeling experiments provide a proof of concept that the four patterns we have highlighted—attested in the recent literature on these games, and mostly unexplained by Gricean competence alone—are natural results of this theory of performance. Because such comprehenders would tend to prefer efficient strategies of interpretation, they would be expected to show higher performance in Simple vs. Complex trials. Because they would revise estimated utilities downward for simpler strategies when they have been ineffective, they would be expected to increase in pragmatic responses over experience in the task. Because they could initially begin with different expectations about the value of social reasoning, their rates of target selection would be expected to vary in association with other tasks that require social reasoning (e.g. the ToM tasks of Mayn and Demberg, 2023b). And crucially, because they would vary in the dynamics of their on-task exploration, their rates of target selection would also be expected to vary in association with any other tasks that require that exploration (e.g. the problem-solving tasks of Mayn and Demberg).

This explanation requires no major changes to the core Gricean hypothesis—i.e. we need not conclude from low performance in Complex trials that comprehenders are incapable of reasoning about the intentions of their interlocutor. Nor do we need to claim that comprehenders are unlikely to perform such reasoning in everyday interactions. Instead, we can see low and variable rates of implicatures in these tasks as the result of a guess that social reasoning may not be worthwhile *in this case*, plus low likelihood of investing the time necessary to revise that guess. And indeed, if a resource-rational comprehender new to this task would realize the worth of pragmatic reasoning within just 24 trials where it is necessary, it is certain that they will have already realized the worth of this reasoning in more common communication scenarios.

We think resource-rational exploration may be relevant in other ways during more naturalistic pragmatic interpretation. It seems to us that interpreting any utterance which is anomalous for typical modes of comprehension, like many utterances which trigger implicatures, likely requires exploration similar to what we model here. Kravtchenko and Demberg (2022) demonstrate one such case, an 'atypicality inference,' where, when speakers explicitly mention an event which should be predictable, comprehenders interpret that the event must have been in fact unexpected. If these cases require identifying that particular inference from a pool of many other hypotheses, it could help explain why Ryzhova et al. (2023) observe an association between rates of atypicality inference and general problem-solving performance, much like we have discussed for the reference game. We plan to extend our model to such cases in future work.

As for reference games, it is clear that more fine-grained data will be needed to validate the approach we take here. We have already begun assembling that data: in Duff et al. (in prep.), we follow up on Mayn and Demberg (2023b), collecting measures of F_{NEG} and persistence in particular, and show that these constructs do in fact control the covariance between the reference game and other problem-solving tasks. We also show that response times follow patterns across conditions and responses which align nicely with ACT-R predictions. We expect that further evaluation of the timecourse predictions of our model will be useful as we continue on the path from theories predicting the likelihood of pragmatic behavior to theories which explain its derivation in real time.

[10] Although there is a longer tradition of algorithmic modeling for referential production, e.g. Dale and Reiter (1995).

Acknowledgments

We're grateful to Sebastian Schuster, Michael Franke, Jakub Dotlačil, Adrian Brasoveanu, and Niels Taatgen for helpful insight and discussion. This project is supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No. 948878).

References

ACT-R Research Group. (2022). *ACT-R tutorial*. Carnegie Mellon University.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Erlbaum.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251.

Brasoveanu, A., & Dotlačil, J. (2020). *Computational cognitive modeling and linguistic theory*. SpringerOpen.

Brochhagen, T. (2021). Brief at the risk of being misunderstood: Consolidating population- and individual-level tendencies. *Computational Brain & Behavior*, *4*, 305–317.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Carstensen, A., Kon, E., & Regier, T. (2014). Testing a rational account of pragmatic reasoning: The case of spatial language. In *Proceedings of CogSci 36* (pp. 2009–2013).

Cheyette, S. J., & Piantadosi, S. T. (2024). Response to difficulty drives variation in IQ test performance. *Open Mind*, *8*, 265–277.

Dale, R., & Reiter, E. (1995). Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, *18*, 233–263.

Degen, J. (2023). The Rational Speech Act framework. *Annual Review of Linguistics*, *9*, 519–540.

Degen, J., & Franke, M. (2012). Optimal reasoning about referential expressions. In *Proceedings of SemDial 16* (pp. 2–11).

Dodell-Feder, D., Lincoln, S. H., Coulson, J. P., & Hooker, C. I. (2013). Using fiction to assess mental state understanding: A new task for assessing theory of mind in adults. *PLoS ONE*, *8*(11), e81279.

Duff, J., Mayn, A., & Demberg, V. (2025). An ACT-R model of resource-rational performance in a pragmatic signaling game: Supplementary materials. https://osf.io/z8459/?view_only=88c11d39450d46a69e645494df7c510e

Duff, J., Mayn, A., & Demberg, V. (in prep.). The role of reinforcement learning in pragmatic reasoning tasks: Modeling and validating the sources of individual differences.

Fairchild, S., & Papafragou, A. (2021). The role of executive function and theory of mind in pragmatic computations. *Cognitive Science*, *45*(2), e12938.

Frank, M. C., Emilsson, A. G., Peloquin, B., Goodman, N. D., & Potts, C. (2016). Rational speech act models of pragmatic reasoning in reference games [Manuscript hosted on PsyArXiV].

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*, 998.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science*, *306*, 1940–1943.

Franke, M. (2011). Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics & Pragmatics*, *4*, 1.

Franke, M., & Degen, J. (2016). Reasoning in reference games: Individual- vs. population-level probabilistic modeling. *PLoS ONE*, *11*(5), e0154854.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Fu, W.-T., & Anderson, J. R. (2006). From recurrent choice to skill learning: A reinforcement-learning model. *Journal of Experimental Psychology: General*, *135*(2), 184–206.

Gonthier, C., Harma, K., & Gavornikova-Baligand, Z. (2024). Development of reasoning performance in Raven's matrices is grounded in the development of effective strategy use. *Journal of Experimental Psychology: General*, *153*(3), 689–705.

Gonthier, C., & Thomassin, N. (2015). Strategy use fully mediates the relationship between working memory capacity and performance on Raven's matrices. *Journal of Experimental Psychology: General*, *144*(5), 916–924.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3). Academic Press.

Hawkins, R. D., Franke, M., Frank, M. C., Goldberg, A. E., Smith, K., Griffiths, T. L., & Goodman, N. D. (2022). From partners to populations: A hierarchical Bayesian account of coordination and convention. *Psychological Review*, *130*(4), 977–1016.

Hawkins, R. D., Gweon, H., & Goodman, N. D. (2021). The division of labor in communication: Speakers help listeners account for asymmetries in visual perspective. *Cognitive Science*, *45*, e12926.

Hayes, T. R., Petrov, A. A., & Sederberg, P. B. (2011). A novel method for analyzing sequential eye movements reveals strategic influence on Raven's Advanced Progressive Matrices. *Journal of Vision*, *11*(10), 10.

Howes, A., Lewis, R. L., & Vera, A. (2009). Rational adaptation under task and processing constraints. *Psychological Review*, *116*(4), 717–751.

Jäger, G. (2010). Game-theoretical pragmatics. In J. F. A. K. van Benthem & A. Ter Meulen (Eds.), *Handbook of logic and language*. Elsevier.

Jarosz, A. F., & Wiley, J. (2012). Why does working memory capacity predict RAPM performance? A possible role of distraction. *Intelligence*, *40*, 427–438.

Kravtchenko, E., & Demberg, V. (2022). Informationally redundant utterances elicit pragmatic inferences. *Cognition*, *225*, 105159.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1.

Mayn, A., & Demberg, V. (2022). Individual differences in a pragmatic reference game. In *Proceedings of CogSci 44* (pp. 3016–3022).

Mayn, A., & Demberg, V. (2023a). High performance on a pragmatic task may not be the result of successful reasoning: On the importance of eliciting participants' reasoning strategies. *Open Mind*, *7*, 156–178.

Mayn, A., & Demberg, V. (2023b). Sources of variability in a pragmatic reference game: Effects of reasoning, memory and perspective-taking [Presented at XPrag 10].

Parikh, P. (1991). Communication and strategic inference. *Linguistics and Philosophy*, *14*, 473–514.

Qing, C., & Franke, M. (2015). Variations on a Bayesian theme: Comparing Bayesian models of referential reasoning. In H. Zeevat & H.-C. Schmitz (Eds.), *Bayesian natural language semantics and pragmatics*. Springer.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scale*. Oxford Psychologists Press.

Rohde, H., Seyfarth, S., Clark, B., Jäger, G., & Kaufmann, S. (2012). Communicating with cost-based implicature: A game-theoretic approach to ambiguity. In *Proceedings of SemDial 16* (pp. 107–116).

Ryzhova, M., Mayn, A., & Demberg, V. (2023). What inferences do people actually make upon encountering informationally redundant utterances? An individual differences study. In *Proceedings of CogSci 45* (pp. 2631–2638).

Sikos, L., Venhuizen, N. J., Drenhaus, H., & Crocker, M. W. (2021). Reevaluating pragmatic reasoning in language games. *PLoS ONE*, *16*(3), e0248388.

Stiller, A. J., Goodman, N. D., & Frank, M. C. (2011). Ad-hoc scalar implicature in adults and children. In *Proceedings of CogSci 33* (pp. 2134–2139).

Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, *11*(2), 176–190.

Stocco, A., Prat, C. S., & Graham, L. K. (2021). Individual differences in reward-based learning predict fluid reasoning abilities. *Cognitive Science*, *45*, e12941.

Storm, B. C., Angello, G., & Bjork, E. L. (2011). Thinking can cause forgetting: Memory dynamics in creative problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1287–1293.

Taatgen, N. A., van Rijn, H., & Anderson, J. (2007). An integrated theory of prospective time interval estimation: The role of cognition, attention, and learning. *Psychological Review*, *114*(3), 577–598.

Trott, S., & Bergen, B. (2019). Individual differences in mentalizing capacity predict indirect request comprehension. *Discourse Processes*, *56*(8), 675–707.

Van Rooy, R. (2004). Signalling games select Horn strategies. *Linguistics and Philosophy*, *27*, 493–527.

Vigneau, F., Caissie, A. F., & Bors, D. A. (2006). Eye-movement analysis demonstrates strategic influences on intelligence. *Intelligence*, *34*, 261–272.